A RE-EXAMINATION OF THE NILE RIVER DATA BASED ON LONG MEMORY AT THE LONG RUN AND THE CYCLICAL FREQUENCIES

L.A. Gil-Alana$^{a,b}$

$^a$ University of Navarra, Pamplona, Spain

ABSTRACT

The issue in this paper is to re-examine the Nile River data by means of using new statistical techniques based on long memory at the long run and the cyclical frequencies. We use a procedure that permits us to consider this type of model, which is based, for the cyclical component, on the Gegenbauer processes. We test for the presence of unit roots with fractional orders of integration at both the zero and the cyclical frequencies. The results show that the root at zero plays a much more important role than the cyclical one, though the latter frequency also displays a component of long memory behaviour.

Keywords: Gegenbauer processes; Fractional cycles; Long memory; Nile River.

Mathematics Subject Classification: 62,10; 62M15; 62P12.

$^b$Correspondence author: Luis A. Gil-Alana
University of Navarra
Faculty of Economics
E-31080 Pamplona, Spain
Phone: 00 34 948 425 625
Fax: 00 34 948 425 626
E-mail: alana@unav.es

The author gratefully acknowledges financial support from the Ministerio de Ciencia y Tecnologia (SEC2005-07657, Spain).
1. Introduction

Data collected from the Nile River have spurred the development of a whole field of mathematics (i.e., fractional Brownian motion and fractional Gaussian noise), along with a field of statistics concerned with the behaviour of long memory time series. In fact, the words of Jarvis (1936) have resulted to be very prophetic: “In spite of all the changing, uncertain, and erroneous factors that must be considered in connection with records of stages of the Nile River, it is believed that they disclose some important information: and there is a fair prospect that they may yield more data with further study and the cummulation of ideas for various students”.

A vast amount of literature has been written about Nile River data. Gathered by Toussoun (1925), the data start in 622 A.D. and end in 1284 A.D., and display the yearly minimal water levels of the river, measured at the Roda Gauge near Cairo. Historically, this data set is of particular interest. Since ancient times, the Nile River has been known for its characteristic long-term behaviour. Long periods of dryness were followed by long periods of yearly returning floods. Floods had the effect of fertilizing the soil so that in flood years the yield of crop was particularly abundant. On the speculative basis, one may find an early qualitative account of this in the Bible (Genesis 41, 29-30): “Seven years of great abundance are coming throughout the land of Egypt, but seven years of famine will follow them”. The analysis of this and several similar time series led to the
discovery of the so-called Hurst effect. In fact, Hurst (1951) was the first in showing that the data possesses long range dependence, though Mandelbrot and van Ness (1968) and Mandelbrot and Wallis (1969) were the pioneers in describing a statistical modelling of this time series in terms of a fractional Gaussian noise model. Beran (1994) also examined this series, using ARFIMA models and concluded that the fractional differencing parameter was around 0.40 with a 95% confidence interval of (0.34, 0.46).

In this article, we model the Nile River data from a different time series perspective and, instead of considering exclusively the component affecting the long run or zero frequency, we also take into account the cyclical structure. For the later component, we make use of the Gegenbauer processes. The structure of the paper is as follows: In Section 2 we briefly describe the literature on modelling trends and cycles in raw time series. In Section 3, we present a procedure that permits us to simultaneously test unit roots with possible fractional orders of integration at zero and the cyclical frequencies. The procedure, due to Robinson (1994a), has several distinguishing features compared with other methods, the most noticeable one being its standard null and local limit distributions. In Section 4, the tests are applied to the Nile River dataset, while Section 5 contains some concluding comments.

2. Modelling trends and cycles in raw time series
It is a well-known stylised fact that many time series contain trends, seasonals as well as cyclical components. However, while the literature on the trend and the seasonal components is quite extensive, little attention has been paid to the cyclical structure. Initially, the trend was explained in terms of deterministic (linear or quadratic) functions of time, which were fitted by regression techniques. Later on, it was shown that, in many series, the trend component changed or evolved over time and stochastic approaches (based on first or second differences of the data) were proposed, especially after the seminal paper of Nelson and Plosser (1982). In that paper, following the work and ideas of Box and Jenkins (1970), and using tests of Fuller (1976) and Dickey and Fuller (1979), they showed that many US macroeconomic series could be specified in terms of unit root processes. Following that work, many test statistics were developed for unit roots, (e.g., Phillips, 1987; Phillips and Perron, 1988; Kwiatkowski et al., 1992; etc.). These processes were later extended to allow for a much more general class of long memory processes, allowing the number of differences to be a fractional value.

For the purpose of the present paper, we define an I(0) process, \{u_t, t = 0, ±1, ...\}, as a covariance stationary process with a spectral density function that is positive and finite at any frequency.\(^1\) In this context, we say that \(x_t\) is I(d) if:

\[
(1 - L)^d x_t = u_t, \quad t = 1, 2, ...., \tag{1}
\]

\(^1\) In other words, \(0 < f(\lambda) < \infty\), where \(f(\lambda)\) is the spectral density function of the process.
with \( x_t = 0, \ t \leq 0, \) where \( L \) is the lag operator \( (Lx_t = x_{t+1}) \) and the unit root case corresponding to \( d = 1 \). If \( d > 0 \) in (1), \( x_t \) is said to be a long memory process because of the strong degree of correlation between observations widely separated in time. If \( d \in (0, 0.5) \), \( x_t \) is covariance stationary, and if \( d \in [0.5, 1) \), \( x_t \) is no longer stationary but is still mean reverting, with the effects of the shocks dying away in the long run. Finally, if \( d \geq 1 \), the series is nonstationary and non-mean-reverting. These processes were introduced by Granger and Joyeux (1980), Granger (1980, 1981) and Hosking (1981), (though earlier work by Adenstedt, 1974, and Taqqu, 1975, show an awareness of its representation), and they were theoretically justified in terms of aggregation by Robinson (1978), Granger (1980), and more recently in terms of the duration of shocks by Parke (1999).

Empirical applications of fractional models like (1) on macroeconomic time series are amongst others the papers of Diebold and Rudebusch (1989), Baillie and Bollerslev (1994), Baillie (1996) and Gil-Alana and Robinson (1997).

Time series may also present a component of long memory at other frequencies rather than zero. The most common case here refers to the seasonal structure, and seasonal long memory processes have been studied in recent years. (See, e.g., Porter-Hudak, 1990, Ray, 1993, Sutcliffe, 1994; Gil-Alana, 2002; etc.). Similarly to the trend and to the seasonal components, deterministic cyclical models have also been studied. Harvey (1985) proposed stochastic cycles, and
Gray et al. (1989, 1994) generalized them to allow for long memory. In particular, they considered processes like:

\[(1 - 2\mu L + L^2)^d x_t = u_t, \quad t = 1,2,\ldots, \quad (2)\]

where \(u_t\) is \(I(0)\). Gray et al. (1989) showed that \(x_t\) in (2) is stationary if \(|\mu| < 1\) and \(d < 0.50\) or if \(|\mu| = 1\) and \(d < 0.25\). They also showed that the polynomial in (2) can be expressed in terms of the Gegenbauer polynomial \(C_{j,d}(\mu)\) such that for all \(d \neq 0\),

\[\frac{(1 - 2\mu L + L^2)^{-d}}{d} = \sum_{j=0}^{\infty} C_{j,d}(\mu)L_j, \quad (3)\]

where

\[C_{j,d}(\mu) = \sum_{k=0}^{\lfloor j/2 \rfloor} \frac{(-1)^k (d)_{j-k} (2\mu)^{-2k}}{k!(j-2k)!}; \quad (d)_j = \frac{\Gamma(d+j)}{\Gamma(d)}, \]

\(\Gamma(x)\) denotes the Gamma function, and a truncation in (3) will be required to make (2) operational. Thus, the process in (3) becomes:

\[x_t = \sum_{j=0}^{t-1} C_{j,d}(\mu)u_{t-j}, \quad t = 1,2,\ldots, \quad (4)\]

and when \(d = 1\), we have:

\[x_t = 2\mu x_{t-1} - x_{t-2} + u_t, \quad t = 1,2,\ldots, \quad (5)\]

which is a cyclical \(I(1)\) process with the periodicity determined by \(\mu\). Tests of (5) based on autoregressive (AR) alternatives were proposed by Ahtola and Tiao (1987). Their tests are embedded in an AR(2) process of form:
which, under the null hypothesis:

\[ H_0: \left| \phi_1 \right| < 2 \text{ and } \phi_2 = -1, \]  

becomes the cyclical I(1) model in (5).

In this article, we combine the trend and the cyclical approaches and consider a model of the form:

\[ (1 - L)^{d_1} (1 - 2\mu L + L^2)^{d_2} x_t = u_t, \quad t = 1, 2, ..., \]  

for given real values \( d_1 \) and \( d_2 \). Thus, \( d_1 \) refers to the stochastic trend, affecting the long run or zero frequency, while \( d_2 \) affects the cyclical part of the series. In the following section, we present a procedure that permits us to test this type of model.

3. The testing procedure

Following Bhargava (1986), Schmidt and Phillips (1992) and others on parameterization of unit-root models, Robinson (1994a) considers the regression model:

\[ y_t = \beta' z_t + x_t \quad t = 1, 2, ..., \]  

where \( y_t \) is a given raw time series; \( z_t \) is a \((k \times 1)\) vector of exogenous variables; \( \beta \) is a \((k \times 1)\) vector of unknown parameters; and the regression errors \( x_t \) are such that:

\[ \rho(L; \theta) x_t = u_t \quad t = 1, 2, ..., \]  

where $\rho$ is a given function which depends on $L$ and the $(p \times 1)$ parameter vector $\theta$, adopting, the form:

$$
\rho (L; \theta) = (1 - L)^{d_1 + \theta_1} (1 - L^4)^{d_s + \theta_s} \prod_{j=2}^{p} (1 - 2 \cos w L + L^2)^{d_j + \theta_j}, \quad (11)
$$

for real given numbers $d_1$, $d_s$, $d_2$, … $d_p$, $\theta_1$, $\theta_2$, …, $\theta_p$, integer $p$, and where $u_t$ is an I(0) process. Under the null hypothesis:

$$
H_0: \quad \theta = 0 \quad (12)
$$

(11) becomes:

$$
\rho (L; \theta = 0) = \rho (L) = (1 - L)^{d_1} (1 - L^4)^{d_s} \prod_{j=2}^{p} (1 - 2 \cos w L + L^2)^{d_j}. \quad (13)
$$

This is a very general specification that permits us to consider different models under the null. For example, if $d_1 = 1$ and $d_s = d_j = 0$ for $j \geq 2$, we have the classical unit-root models (Dickey and Fuller, 1979, Phillips and Perron, 1990, Kwiatkowski et al., 1992, etc.), and, if $d_1$ is a real value, the fractional models examined in Diebold and Rudebusch (1989), Baillie (1996) and others. Similarly, if $d_s = 1$ and $d_j = 0$ for all $j$, we have the seasonal unit root model (Dickey, Hasza and Fuller, 1984, Hylleberg et al., 1990, etc.) and if $d_s$ is real, the seasonal fractional model analysed in Porter-Hudak (1990). Finally, if $d_1 = d_s = 0$, we have the k-factor Gegenbauer processes studied in Ferrara and Guegan (2001).

In this article we are interested in the long run and the cyclical structures of the series and thus, we take $d_s = 0$ and $p = 2$. In such a case (11) can be re-written as:
A RE-EXAMINATION OF THE NILE RIVER DATA                         517

\[ \rho(L; \theta) = (1 - L)^{d_1 + d_2} (1 - 2 \cos \omega L + L^2)^{d_2 + d_3} \]  
(14)

and similarly (13), \( \rho(L) = (1 - L)^{d_1} (1 - 2 \cos \omega L + L^2)^{d_2} \). Here, \( d_1 \) represents the degree of integration at the zero frequency, i.e., referring to the stochastic trend, while \( d_2 \) affects the cyclical component of the series.

We next describe the test statistic. We observe \( \{(y_t, z_t), t = 1, 2, \ldots n\} \), and suppose that the I(0) \( u_t \) in (10) have spectral density given by:

\[ f(\lambda; \tau) = \frac{\sigma^2}{2\pi} g(\lambda; \tau), \quad -\pi < \lambda \leq \pi, \]

where \( g \) is a function of known form which depends on frequency \( \lambda \) and the unknown \((q \times 1)\) vector \( \tau \). Based on \( H_0 \) (12), the residuals in (9); (10) and (14) are

\[ \hat{u}_t = (1 - L)^{d_1} (1 - 2 \cos \omega L + L^2)^{d_2} y_i - \hat{\beta} s_i, \]  
(15)

\[ \hat{\beta} = \left( \sum_{t=1}^{n} s_t s_t^t \right)^{-1} \sum_{t=1}^{n} s_t (1 - L)^{d_1} (1 - 2 \cos \omega L + L^2)^{d_2} y_i, \]

\[ s_i = (1 - L)^{d_1} (1 - 2 \cos \omega L + L^2)^{d_2} z_i. \]

Unless \( g \) is a completely known function (e.g., \( g \equiv 1 \), as when \( u_t \) is white noise), we have to estimate the nuisance parameter \( \tau \), for example by \( \hat{\tau} = \arg \min_{\tau \in \mathbb{T}} \sigma^2(\tau) \), where \( \mathbb{T} \) is a suitable subset of \( \mathbb{R}^q \) Euclidean space, and

\[ \sigma^2(\tau) = \frac{2\pi}{n} \sum_{s=1}^{n-1} \frac{1}{g(\lambda_s; \tau)} I_{\hat{g}}(\lambda_s), \]

with
Thus, the tests are purely parametric, requiring specific modelling assumptions for the short memory specification of \( u_t \). If \( u_t \) is, for example, an AR process of form 
\[
\phi(L)u_t = \epsilon_t,
\]
then \( g = |\phi(e^{i\lambda})|^2 \), with \( \sigma^2 = V(\epsilon_t) \), so that the AR coefficients are a function of \( \tau \).

The test statistic, which is derived via Lagrange Multiplier (LM) principle, adopts the form:

\[
\hat{R} = \frac{n}{\hat{\sigma}^4} \hat{A}^T \hat{A}^{-1} \hat{a},
\]

where \( n \) is the sample size, and

\[
\hat{a} = -\frac{2\pi}{n} \sum_s \psi(\lambda_s) g(\lambda_s, \hat{\tau})^{-1} I(\lambda_s);
\]

\[
\hat{\sigma}^2 = \sigma^2(\hat{\tau}) = \frac{2\pi}{n} \sum_{s=1}^{n-1} g(\lambda_s, \hat{\tau})^{-1} I(\lambda_s),
\]

\[
\hat{A} = \frac{2}{n} \left( \sum_s \psi(\lambda_s) \psi(\lambda_s)' - \sum_s \psi(\lambda_s) \hat{\epsilon}(\lambda_s) ' \left( \sum_s \hat{\epsilon}(\lambda_s) \hat{\epsilon}(\lambda_s)' \right)^{-1} \sum_s \hat{\epsilon}(\lambda_s) \psi(\lambda_s)' \right)
\]

\[
\psi(\lambda_s)' = \begin{bmatrix} \psi_1(\lambda_s) & \psi_2(\lambda_s) \end{bmatrix}; \quad \hat{\epsilon}(\lambda_s) = \frac{\partial}{\partial \tau} \log g(\lambda_s, \hat{\tau});
\]

\[
\psi_1(\lambda_s) = \log \left| 2 \sin \frac{\lambda_s}{2} \right|; \quad \psi_2(\lambda_s) = \log \left| 2 \left( \cos \lambda_s - \cos \omega \right) \right|;
\]
and the summation on * in the above expressions is over the bounded frequencies in the spectrum.

Based on $H_0$ (12), Robinson (1994a) established that, under certain regularity conditions:

$$
\hat{R} \to_d \chi^2_2, \quad as \quad n \to \infty.
$$

Thus, unlike other procedures, we are in a classical large-sample testing situation by reasons described in Robinson (1994a), who also showed that the tests are efficient in the Pitman sense against local departures from the null. A test of (12) will reject $H_0$ against the alternative $H_a$: $\theta \neq 0$ if $\hat{R} > \chi^2_{2,\alpha}$, where $\text{Prob}(\chi^2_{2,\alpha} > \chi^2_{2,\alpha}) = \alpha$. This version of Robinson’s (1994a) tests was examined in Gil-Alana (2001) and its performance in the context of unit root cycles was compared with Ahtola and Tiao’s (1987) tests, the results showing that Robinson’s (1994a) tests outperform Ahtola and Tiao (1987) in a number of cases.

4. The cyclical structure of the Nile River data

The time series data analysed in this section correspond to the yearly minimum water levels at the Roda Gauge (622-1281 A.D.) obtained by Tousson (1925) and they can also be found in Beran (1994, pp.237-239).

---

2 These conditions are very mild, regarding technical assumptions to be satisfied by $\psi_1(\lambda)$ and $\psi_2(\lambda)$.

3 Tousson (1925) collected in fact data for the Nile River minima up until 1921, with 622-1284 being the longest stretch of data without a gap.
Figure 1 displays plots of the original series and its first differences, along with their corresponding correlograms and periodograms. Starting with the original series, it clearly exhibits a long-term behaviour that might partly give an "explanation" of the seven "good" years and seven "bad" years described in Genesis. There are long periods where the maximal level tends to stay high. On the other hand, there are long periods with low levels. Overall, the series looks stationary, though a visual inspection at the correlogram shows that there are significant values even at some lags relatively far away from zero. Also, the periodogram presents a large value at the smallest frequency, which may be an indication that fractional differences are required to achieve I(0) stationarity. This is corroborated by the plots of the first differenced data. In fact, the correlogram and the periodogram suggest that the series is now over differenced at the zero frequency.

The first thing we do is to compute some statistics in relation with the zero frequency. First, we perform a version of the tests of Robinson (1994a) to check for the presence of roots at such a frequency. We consider a model given by (9) and (10), with \( z_t = (1,t)^\prime \), \( t \geq 1 \), and \( \rho(L; \theta) = (1 - L)^{d+\theta} \), i.e., under \( H_0 \) (12), we test the model:

\[
y_t = \beta_0 + \beta_1 t + x_t, \quad t = 1, 2, \ldots
\]

(18)

\[
(1 - L)^d x_t = u_t, \quad t = 1, 2, \ldots,
\]

(19)
A RE-EXAMINATION OF THE NILE RIVER DATA

for values $d_0 = 0, (0.01), 2$, and different types of disturbances. Initially, we assume that $\beta_0 = \beta_1 = 0$ a priori, (i.e., we do not include any regressors in the undifferenced regression model (18)), though we also consider the cases of an intercept, ($\beta_0$ unknown and $\beta_1 = 0$ a priori), and an intercept and a linear time trend, ($\beta_0$ and $\beta_1$ unknown). Thus, for example, if $u_t$ is white noise and $d_o = 1$, the differences $(1 - L)y_t$ behave, for $t > 1$, like a random walk when $\beta_1 = 0$, and a random walk with a drift when $\beta_0, \beta_1 \neq 0$. The test statistic, denoted here by $\hat{R}_o$, adopts then a similar form to (16), the only difference being that $\psi(\lambda_s)$ is now a scalar equal to $\psi_1(\lambda_s)$, and the residuals in (15):

$$\hat{u}_t = (1 - L)^d y_t - \hat{\beta} s_t, \quad \hat{\beta} = \left( \sum_{t=1}^{n} s_t s_t' \right)^{-1} \sum_{t=1}^{n} s_t (1 - L)^d y_t; \quad s_t = (1 - L)^d z_t.$$

Besides, for this particular specification of $\rho(L; \theta)$ in (10), $p = \dim(\theta) = 1$ and thus, we can consider one-sided versions of the tests. The test statistic reported across Table 1 corresponds to the one-sided statistic $\hat{r} (= \sqrt{\hat{R}_o})$. Thus, a one-sided 100$\alpha$% level test of $H_o$ (12) against the alternative: $H_a$: $\theta > 0$ ($\theta < 0$) will reject $H_o$ (12) if $\hat{r} > z_{\alpha}$ ($\hat{r} < -z_{\alpha}$), where the probability that a standard normal variate exceeds $z_{\alpha}$ is $\alpha$. However, instead of present the whole battery of results for each $d$ and each type of regressors, we report in Table 1 the confidence intervals of those values of $d$ where $H_o$ (12) cannot be rejected at the 5% significance level.
We perform the tests assuming that the disturbances are both white noise and weakly autocorrelated. Starting with the case of white noise $u_t$, we observe that if we do not include regressors, the values of $d$ range between 0.68 and 0.75. However, including an intercept and/or a linear time trend, they are slightly smaller, oscillating between 0.35 and 0.45. Next, we consider the case of AR(1) and AR(2) disturbances. Here, if we do not include regressors, there is a lack of monotonicity in the value of the test statistic with respect to $d$. This monotonicity is a characteristic of any reasonable statistic, given correct specification and adequate sample size, because it is a one-sided statistic. Thus, for example, if $H_0$ (12) is rejected for $d = 1$ against alternatives of form: $\theta > 0$, an even more significant result in this direction should be expected when $d = 0.90$ or 0.80 are tested. The lack of monotonicity may be an indication of model misspecification: frequently, misspecification inflates both numerator and denominator of $\hat{r}$ (and $\hat{R}$) to varying degrees, and affects $\hat{r}$ in a complicated way. (Gil-Alana and Robinson, 1997). Thus, computing $\hat{r}$ for a range of values of $d$ is useful in revealing possible misspecification, though monotonicity is by no means necessarily strong evidence of correct specification. In order to solve this problem, we also employ a non-parametric approach to model the I(0) $u_t$, which is due to Bloomfield (1973). In his model, $u_t$ is exclusively described by the spectral density function, and the function $g$ below (14) is given by:
\[ g(\lambda, \tau) = \exp\left(2 \sum_{l=0}^{k} \tau_l \cos(\lambda_l l)\right). \]  

Like the stationary AR model, this has exponentially decaying autocorrelations. Formulae for Newton-type iteration for estimating the \( \tau_l \) are very simple (involving no matrix inversion), updating formulae when \( k \) is increased are also simple, and we can replace \( \hat{A} \) below (16) by the population quantity:

\[ \sum_{l=k+1}^{\infty} l^{-2} = \frac{\pi^2}{6} - \sum_{l=1}^{k} l^{-2}, \]

which indeed is constant with respect to the \( \tau_l \) (unlike what happens in the AR case). The results based on Bloomfield (1973) disturbances are also displayed in Table 1. They are very similar to those obtained for the case of white noise \( u_t \), with \( d \) oscillating between 0.66 and 0.83 for the case of no regressors, and between 0.36 and 0.56 with deterministic trends. The differences observed in the results between the cases of no regressors and those based on an intercept with or without a time trend are significant from a statistical viewpoint. Thus, in the first case, \( d \) is higher than 0.5, implying nonstationarity, while in the cases of an intercept and/or a linear time trend \( d \) appears to be smaller than 0.5 and thus implying stationarity. We are more inclined to believed that the former approach better approximates to the data, the reason being that the latter model (i.e., with no regressors) assumes a zero-mean value for the series while they are all necessarily positive values.
Still at the zero frequency, it may also be of interest, to examine \( d \), independently of the way of modelling the I(0) disturbances. Here, we use a semiparametric Whittle procedure (Robinson, 1995a) that we are now to describe.

The method of Robinson (1995a) is basically a ‘Whittle estimate’ in the frequency domain, given a band of frequencies that degenerates to zero. The estimate is implicitly defined by:

\[
\hat{d} = \arg \min_d \left( \log \frac{C(d)}{2d} - 2d \sum_{s=1}^{m} \log \lambda_s \right),
\]

where \( C(d) = \frac{1}{m} \sum_{s=1}^{m} I_x(\lambda_s) \lambda_s^{2d} \), \( \lambda_s = \frac{2\pi s}{n} \), \( m \to n \to 0 \),

\[
\begin{align*}
\lambda_s &= \frac{2\pi s}{n}, \\
m &= \frac{\pi}{d_o}.
\end{align*}
\]

Under finiteness of the fourth moment and other mild conditions, Robinson (1995a) proved that:

\[
\sqrt{m} (\hat{d} - d_o) \to_d N(0, 1/4) \text{ as } n \to \infty,
\]

where \( d_o \) is the true value of \( d \) and with the only additional requirement that \( m \to \infty \) slower than \( n \). Robinson (1995a) shows that \( m \) should be smaller than \( n/2 \). A multivariate extension of this estimation procedure can be found in Lobato (1999). There also exist other semiparametric procedures for estimating the fractional differencing parameter, for example, the log-periodogram regression estimate
(LPE), initially proposed by Geweke and Porter-Hudak (1983) and modified later by Künsch (1986) and Robinson (1995b) and the averaged periodogram estimate (APE) of Robinson (1994b). However, we have decided to use in this article the Whittle approach because of its computational simplicity: Using the Whittle method we do not need to employ any additional user-chosen numbers in the estimation (as is the case with the modified LPE and the APE). Also, Gaussianity is not required in order to obtain the asymptotic normal distribution in (22), this method being more efficient than the LPE.

Figure 2 reports the results based on Robinson (1995a), i.e., \( \hat{\Delta} \) given by (21) for a range of values of m from 50 to (n/2)-1.\(^5\) We see that for all m, the values are between 0.35 and 0.50, implying again long memory and stationarity with respect to the long run or zero frequency.

The use of the above approaches to investigate the long run behaviour of the time series consists of testing a parametric model and estimating a semiparametric one, relying on the long run implications of the estimated model. The primary advantage is the precision gained by focusing the information in the series through the parameter estimates. A drawback is that the parameter estimates are sensitive to the class of models considered, and may be misleading because of

---

\(^4\) Velasco (1999a, b) has recently showed that the fractionally differencing parameter can also be consistently semiparametrically estimated in nonstationary contexts by means of tapering. See also Phillips and Shimotsu (2004, 2005).

\(^5\) Some attempts to calculate the optimal bandwidth numbers have been examined in Delgado and Robinson (1996) and Robinson and Henry (1996). However, in the case of the Whittle estimator,
misspecification. However, the possibility of misspecification with parametric (or even semiparametric) models can never be settled conclusively, and the problem can be addressed by considering a larger class of models. This is the approach used here and for this purpose, we employ a larger version of the tests of Robinson (1994a) that permits us to simultaneously consider roots at zero and the cyclical frequencies.

We consider the model given by (10) and (18), with $\rho(L; \theta)$ as in (14). Thus, under $H_0$ (12), the model becomes:

$$y_t = \beta_0 + \beta_1 t + x_t, \quad t = 1, 2, ...$$  \hspace{1cm} (22)

$$(1 - L)^{d_1} (1 - 2 \cos \omega L + L^2)^{d_2} x_t = u_t, \quad t = 1, 2, ..., (23)$$

and if $d_2 = 0$, the model reduces to the case previously studied of long memory exclusively at the long run or zero frequency. We assume that $w = w_r = 2\pi/r$, $r$ indicating the number of time periods per cycle.

We computed the statistic $\hat{R}$ given by (16) for values of $d_1$ and $d_2 = 0.05$, $(0.05)$, 2, and $r = 2, n/2,^6$ assuming that $u_t$ is white noise. However, instead of presenting the results for all values of $r, d_1$ and $d_2$, we just report across Table 2 those cases where we observe non-rejection values at the 95% significance level. We see that if we do not include regressors, the null is always rejected, and including an intercept, and an intercept with a linear trend, the results are similar

---

6 The use of optimal values has not been theoretically justified. Other authors, such as Lobato and Savin (1998) use an interval of values for $m$. 
in both cases, suggesting that the presence of a time trend is not required when modelling this series.\(^7\) We also observe that the values of \(r\) where \(H_0\) (12) cannot be rejected are constrained between 4 and 10, which is consistent with the prophetic words in the Bible that the cycles occur at approximately seven years. Also, \(d_1\) is in all cases higher than \(d_2\). Thus, the non-rejection values of \(d_1\) are 0.35 and 0.40, while for \(d_2\) they are 0.05 and 0.10. In the light of this, it seems clear that the root at the zero frequency plays a much more important role than the one at the cyclical frequency, though the latter also appears to have a component of long memory behaviour.

In order to be more precise about the non-rejection values of \(d_1\) and \(d_2\), we re-computed the tests but this time for a shorter grid, with \(d_1, d_2 = 0.01, (0.01), 2\). Figure 3 displays the regions of \((d_1, d_2)\)-values where \(H_0\) (12) cannot be rejected at the 5% level, including an intercept, and with \(r = 5, 6, 7\) and 8. We observe that the values are similar for the four cases. Thus, the values of \(d_1\) oscillates between 0.3 and 0.5, while \(d_2\) is constrained between 0 and 0.15, implying that the series might be stationary with respect to both the zero and the cyclical frequency. Note that we have deliberately excluded in the computation the case of \(d_2 = 0\), the reason being that the model reduces then to the previous case of long memory at the long run or zero frequency. However, it should be noted that the null

\[^6\text{Note that if } r = 1, \text{ the cyclical part reduces to an I(d) process, with the singularity occurring exclusively at the long run or zero frequency.}\]
hypothesis cannot be rejected in some cases here with $d_1 \in [0.35, 0.50]$, which is consistent with the previous results, though these hypothesis were generally less clearly non-rejected than with $d_2 > 0$.\footnote{In fact, the coefficients corresponding to the time trend were insignificantly different from zero in all cases where the null could not be rejected. Note that the tests are based on the null differenced model, which is short memory, and thus, standard t-tests apply.} The tests were also performed allowing autocorrelated disturbances and, though we do not display the results in the paper, they were very similar to those reported here, with values of $d_1$ and $d_2$ smaller than 0.5 and higher for the long run or zero frequency.

5. Concluding comments

In this article we have examined the time series behaviour of data collected from the Nile River by means of new statistical techniques based on long memory processes. We have used a procedure due to Robinson (1994a) that permits us to consider unit roots with fractional orders of integration not only at zero but also at the cyclical frequencies. The tests have standard null and local limit distributions and are easy to implement in raw time series.

These data have been examined by many authors, most of them concentrating exclusively on the long run or zero frequency. Beran (1994), for example, concluded that the series could be specified in terms of an ARFIMA model with $d$ around 0.40. We initially performed a version of Robinson’s (1994a) tests that permits to examine the root at such a frequency. The results

\footnote{By "less clearly non-rejected" we mean that the test statistic was closer to the critical value.}
show that $d$ is higher than 0.5 if no regressors are included in the model. However, the values of $d$ are constrained between 0.35 and 0.45 if an intercept and/or a linear time trend is included whether the disturbances are or not autocorrelated. Similar conclusions were obtained when using a semiparametric Whittle procedure (Robinson, 1995a), a result that is completely in line with Beran (1994).

Next, we perform a procedure for simultaneously testing for the presence of roots at zero and the cyclical components. For the latter frequency, the model is based on the Gegenbauer processes. Here the procedure provides us with some interesting results. Thus, the periodicity of the series seems to be constrained between 5 and 9 years, which is consistent with the earlier discussion that cycles have a length of approximately seven years. Also, the orders of integration seem to fluctuate between 0.30 and 0.50 at the long run frequency and between 0 and 0.15 for the cyclical one, implying that the series is stationary with respect to both the zero and the cyclical frequencies. Moreover, the fact that $d_1$ is in all cases higher than $d_2$ suggests that the root at zero plays a much more important role than the one corresponding to the cyclical frequency, though the latter also seems to present a component of long memory behaviour.

It would also be worthwhile proceeding to get point estimates for the fractional differencing parameters in this context of trends and cyclical models. For the trending component the literature is huge (see, e.g., Fox and Taqqu, 1986;
Dahlhaus, 1989; Sowell, 1992; etc.). For the cyclical part, some attempts have been made by Arteche and Robinson (2000) and Arteche (2002). However, the goal of this paper is to show that a fractional model with the roots simultaneously occurring at both the zero and the cyclical frequencies can be a credible alternative when modelling many time series and, in that respect, the results presented in this paper leads us to some unambiguous conclusions, with the periodicity constrained between 4 and 9 years and the orders of integration higher than 0 but smaller than 0.5, implying stationarity and long memory in relation with both the zero and the cyclical components.

References


[48] P.M. Robinson, Log-periodogram regression of time series with long range
[49] P.M. Robinson, Gaussian semiparametric estimation of long range
[50] P.M. Robinson and M. Henry, Bandwidth choice in Gaussian
semiparametric estimation of long-range dependence. P.M. Robinson and M.
Rosenblatt eds. Athens Conference on Applied Probability in Time Series
[51] P. Schmidt, and P.C.B. Phillips, LM tests for a unit root in the presence of
257-287.
[52] F. Sowell, Maximum likelihood estimation of stationary univariate
165-188.
[53] A. Sutcliffe, Time series forecasting using fractional differencing, Journal
[54] M.S. Taqqu, Weak convergence to fractional motion and to the Rosenblatt
[55] O. Toussoun, Memoire sur l’histoire du Nil, In Memoires a l’Institut
d’Egipite, 18,(1925), 366--404.
A RE-EXAMINATION OF THE NILE RIVER DATA


FIGURE 1

Nile River data and its first differences, with their corresponding correlograms and periodograms

<table>
<thead>
<tr>
<th>Original time series</th>
<th>First differences</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Correlogram original series" /></td>
<td><img src="image" alt="Correlogram first differences" /></td>
</tr>
<tr>
<td><img src="image" alt="Periodogram original series" /></td>
<td><img src="image" alt="Periodogram first differences" /></td>
</tr>
</tbody>
</table>
The large sample standard error under the null hypothesis of no autocorrelation is $1/\sqrt{T}$ or roughly 0.038.

### TABLE 1

<table>
<thead>
<tr>
<th></th>
<th>No regressors</th>
<th>An intercept</th>
<th>A linear time</th>
</tr>
</thead>
<tbody>
<tr>
<td>White noise</td>
<td>[0.68 - 0.75]</td>
<td>[0.36 - 0.45]</td>
<td>[0.35 - 0.45]</td>
</tr>
<tr>
<td>AR(1)</td>
<td>Lack of monotonicity</td>
<td>[0.31 - 0.44]</td>
<td>[0.29 - 0.43]</td>
</tr>
<tr>
<td>AR(2)</td>
<td>Lack of monotonicity</td>
<td>[0.31 - 0.52]</td>
<td>[0.28 - 0.51]</td>
</tr>
<tr>
<td>Bloomfield (1)</td>
<td>[0.73 - 0.86]</td>
<td>[0.31 - 0.44]</td>
<td>[0.29 - 0.44]</td>
</tr>
<tr>
<td>Bloomfield (2)</td>
<td>[0.66 - 0.83]</td>
<td>[0.36 - 0.52]</td>
<td>[0.41 - 0.56]</td>
</tr>
</tbody>
</table>

### FIGURE 2

Estimates of $d$ based on the Whittle semiparametric method (Robinson, 1995a)
<table>
<thead>
<tr>
<th>r</th>
<th>$d_1$</th>
<th>$d_2$</th>
<th>No regressors</th>
<th>An intercept</th>
<th>A linear trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.40</td>
<td>0.05</td>
<td>774.74</td>
<td>0.693</td>
<td>0.542</td>
</tr>
<tr>
<td>4</td>
<td>0.40</td>
<td>0.10</td>
<td>872.03</td>
<td>5.365</td>
<td>4.935</td>
</tr>
<tr>
<td>4</td>
<td>0.45</td>
<td>0.05</td>
<td>502.54</td>
<td>1.384</td>
<td>1.549</td>
</tr>
<tr>
<td>4</td>
<td>0.45</td>
<td>0.10</td>
<td>580.10</td>
<td>2.112</td>
<td>2.161</td>
</tr>
<tr>
<td>4</td>
<td>0.50</td>
<td>0.10</td>
<td>354.95</td>
<td>4.700</td>
<td>4.841</td>
</tr>
<tr>
<td>5</td>
<td>0.35</td>
<td>0.05</td>
<td>876.92</td>
<td>4.606</td>
<td>3.710</td>
</tr>
<tr>
<td>5</td>
<td>0.40</td>
<td>0.05</td>
<td>609.05</td>
<td>0.584</td>
<td>0.538</td>
</tr>
<tr>
<td>5</td>
<td>0.40</td>
<td>0.10</td>
<td>649.43</td>
<td>4.156</td>
<td>3.972</td>
</tr>
<tr>
<td>5</td>
<td>0.45</td>
<td>0.05</td>
<td>388.81</td>
<td>2.234</td>
<td>2.394</td>
</tr>
<tr>
<td>5</td>
<td>0.45</td>
<td>0.10</td>
<td>420.30</td>
<td>4.759</td>
<td>4.858</td>
</tr>
<tr>
<td>6</td>
<td>0.35</td>
<td>0.05</td>
<td>815.56</td>
<td>4.092</td>
<td>3.282</td>
</tr>
<tr>
<td>6</td>
<td>0.40</td>
<td>0.05</td>
<td>560.13</td>
<td>1.598</td>
<td>1.552</td>
</tr>
<tr>
<td>6</td>
<td>0.40</td>
<td>0.10</td>
<td>565.77</td>
<td>6.037</td>
<td>5.944</td>
</tr>
<tr>
<td>6</td>
<td>0.45</td>
<td>0.05</td>
<td>353.13</td>
<td>3.853</td>
<td>3.989</td>
</tr>
<tr>
<td>7</td>
<td>0.35</td>
<td>0.05</td>
<td>799.22</td>
<td>1.596</td>
<td>1.077</td>
</tr>
<tr>
<td>7</td>
<td>0.35</td>
<td>0.10</td>
<td>778.71</td>
<td>9.144</td>
<td>5.410</td>
</tr>
<tr>
<td>7</td>
<td>0.40</td>
<td>0.05</td>
<td>541.74</td>
<td>1.077</td>
<td>1.117</td>
</tr>
<tr>
<td>7</td>
<td>0.45</td>
<td>0.05</td>
<td>3351.76</td>
<td>4.680</td>
<td>4.879</td>
</tr>
<tr>
<td>8</td>
<td>0.35</td>
<td>0.05</td>
<td>815.65</td>
<td>2.153</td>
<td>1.621</td>
</tr>
<tr>
<td>8</td>
<td>0.40</td>
<td>0.05</td>
<td>550.31</td>
<td>2.296</td>
<td>2.375</td>
</tr>
<tr>
<td>9</td>
<td>0.35</td>
<td>0.05</td>
<td>840.37</td>
<td>2.372</td>
<td>1.882</td>
</tr>
<tr>
<td>9</td>
<td>0.40</td>
<td>0.05</td>
<td>564.32</td>
<td>3.338</td>
<td>3.440</td>
</tr>
<tr>
<td>10</td>
<td>0.35</td>
<td>0.05</td>
<td>864.31</td>
<td>2.064</td>
<td>1.672</td>
</tr>
<tr>
<td>10</td>
<td>0.40</td>
<td>0.05</td>
<td>377.35</td>
<td>3.944</td>
<td>4.102</td>
</tr>
</tbody>
</table>

In bold: Non-rejection values of the null hypothesis (12) at the 5% significance level.
FIGURE 3

Values of $d_1$ and $d_2$ where $H_0$ (12) cannot be rejected at the 5% significance level

$r = 5$

$r = 6$

$r = 7$

$r = 8$
The horizontal axe refers to $d_1$, the order of integration at the zero frequency while the vertical one is $d_2$, the order of integration at the cyclical frequency.

**Received: January 21, 2006**