ESTIMATION OF THE FRACTIONAL DIFFERENCE PARAMETER IN THE ARIMA(p, d, q) MODEL USING THE SMOOTHED PERIODOGRAM

BY VALDERIO A. REISEN

UMIST

First version received June 1993

Abstract. In recent work on time series analysis considerable interest has been focused on series having the property of *long memory*. Long memory is a characteristic of time series in which the dependence between distant observations is not negligible. The model that has been most frequently studied, which in some situations shows properties of long memory, is based on the autoregressive integrated moving-average ARIMA(p, d, q) process. Hosking (Fractional differencing, *Biometrika* 68 (1) (1981), 165–76) generalized this model by permitting the *degree of differencing d* to take fractional values. He then demonstrated that for d in the range $0 < d < 0.5$ the process is stationary and possesses the long memory property. Our study is based on the ARIMA(p, d, q) model when d takes any real non-integer value in the interval $(-0.5, 0.5)$. The main aim of our study is to examine methods for estimating the parameters of this model. For estimating d we suggest an estimator based on the smoothed periodogram. Using an empirical approach we compare this estimator with other which are well known in the literature of long memory models, e.g. the raw periodogram regression method and the Hurst coefficient method.

Keywords. General fractional differenced white noise process; degree of differencing; long memory; periodogram regression; smoothed periodogram regression; Hurst coefficient.

1. INTRODUCTION

The autoregressive integrated moving-average ARIMA(p, d, q) process shows the long memory property when the parameter d takes certain non-integer values. 'Long memory' is a characteristic of a time series in which the dependence between distant observations is not negligible, and series with this property may be characterized by

(i) the autocorrelation function not being absolutely summable and
(ii) the spectral density function becoming unbounded as the frequency tends to zero.

Hosking (1981) demonstrated that the ARIMA(p, d, q) process with $d \in (-0.5, 0.5)$ is stationary and invertible, and for $0 < d < 0.5$ has the long memory property.

The main purpose of this study is to examine methods for estimating the parameter d in the model above. The technique of the regression method based on the spectral density of the process was first studied by Geweke and

© 1994 Basil Blackwell Ltd., 108 Cowley Road, Oxford OX4 1JF, UK and 238 Main Street, Cambridge, MA 02142, USA.

Porter-Hudak (1983). They proposed a method for estimating d using a regression model based on the periodogram. Using the asymptotic normal distribution of the smoothed periodogram we now propose an estimate of d based on a regression of the smoothed periodogram. We will show by simulation studies that the later method generally results in an estimate with a smaller bias and variance. We derive theoretical results for both regression methods. However, the theoretical results hold only when $d < 0$ but we will show that our simulation studies support the validity of these methods even in the case of positive d. Using both functions, i.e. the periodogram and the smoothed periodogram, in the regression method, we also present results for determining whether or not a time series has a long memory property. Another approach to the estimation of d studied here is the Hurst coefficient method, first proposed by Hurst (1951, 1956) and later reviewed by McLeod and Hipel (1978). The outline of this paper is thus as follows. In Section 2 we summarize the definitions and properties of the general fractional ARIMA(p, d, q) model. In Section 3 we present theoretical results underlying the methods of estimating d described above. In Section 4 we present simulated results comparing the three methods

2. THE ARIMA(p, d, q) MODEL: DEFINITIONS AND PROPERTIES

As noted above, we allow the parameter d in the ARIMA(p, d, q) model to take a non-integer real value. We now present the following definitions.

Let $\{\varepsilon_t\}$ be a white noise process with $E(\varepsilon_t) = 0$, $\sigma_\varepsilon^2 > 0$, and let B be the back-shift operator, i.e. $BX_t = X_{t-1}$. Let $\Phi(B)$ and $\Theta(B)$ be polynomials of orders p and q respectively where $\Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\Theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$.

If $\{X_t\}$ is a linear process satisfying

$$\Phi(B)(1 - B)^d X_t = \Theta(B)\varepsilon_t \qquad \text{for } d \in (-0.5, 0.5) \tag{1}$$

then $\{X_t\}$ is called a general fractional differenced white noise (GFDWN) process; d is called the degree of differencing. The polynomials $\Phi(B)$ and $\Theta(B)$ have all roots outside the unit circle and share no common roots.

When $\Phi(B) \equiv \Theta(B) \equiv 1$, $\{X_t\}$ is called a fractional differenced white noise (FDWN) process and is represented by $(1 - B)^d X_t = \varepsilon_t$. These definitions are a natural extension of the terminology of Box and Jenkins (1976) to the case of a non-integral parameter d.

Note that the term $(1 - B)^d$ in (1), for $d \in R$, is defined by a binomial expansion

$$(1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k = 1 - dB - \frac{d}{2!}(1 - d)B^2 - \cdots.$$

The GFDWN for fractional values of d was introduced by Hosking (1981) where he described the basic properties of this model; we summarize some of these properties below.

Let $d \in (-0.5, 0.5)$. Then the following hold.

(i) The GFDWN process is stationary and invertible.

(ii) The spectral density of the process is given by $f(w) = f_u(w)\{2\sin(w/2)\}^{-2d}$ where $f_u(w)$ is the spectral density of the ARIMA(p, q) process $U_t = (1 - B)^d X_t$. Thus,

$$f(w) = \sigma_\varepsilon^2 \frac{|\Theta(e^{-iw})|^2}{2\pi|\Phi(e^{-iw})|^2}\left\{2\sin\left(\frac{w}{2}\right)\right\}^{-2d} \qquad w \in [-\pi, \pi]. \tag{2}$$

As $w \to 0$, $\lim\{w^{2d}f(w)\}$ exists and is finite.

(iii) Let ρ_k^x be the autocorrelation function of $\{X_t\}$. Then, as $k \to \infty$, $\lim(k^{1-2d}\rho_k^x)$ exists and is finite.

The properties (ii) and (iii) show the the GFDWN process has the 'long memory' property for positive values of d in the specified range, i.e. the spectral density is unbounded near zero frequency and the autocorrelation function is not absolutely summable. Many authors have used the GFDWN process in their work, including Granger and Joyeux (1980), Andel (1986), Parzen (1986), Hassler (1993) and Geweke and Porter-Hudak (1983).

3. ESTIMATES OF THE FRACTIONAL PARAMETER d

We now describe the three methods for estimating d referred to in the preceding section. (These methods will be compared by simulation studies in the next section.) We first introduce the estimate of d using both regression methods. Consider the set of harmonic frequencies $w_j = 2\pi j/n$, $j = 0, 1, \ldots, [n/2]$, where n is the sample size. As in the previous section let $\{X_t\}$ be an ARIMA(p, d, q) process with $d \in (-0.5, 0.5)$. Taking the logarithm of the spectral density $f(w)$ given by (2) we have

$$\ln\{f(w_j)\} = \ln\{f_u(w_j)\} - d\ln\left\{2\sin\left(\frac{w_j}{2}\right)\right\}^2$$

which may be written in the alternative form

$$\ln\{f(w_j)\} = \ln(f_u(0)\} - d\ln\left\{2\sin\left(\frac{w_j}{2}\right)\right\}^2 + \ln\left\{\frac{f_u(w_j)}{f_u(0)}\right\}. \tag{3}$$

From equation (3) we may construct two regression equations as follows.

3.1. *The estimator of d using the periodogram*

Given a sample of observations on X_t, say $X_1, X_2, \ldots, X_n$, the periodogram is defined by

$$I_x(w) = \frac{1}{2\pi}\left\{R(0) + 2\sum_{s=1}^{n-1} R(s)\cos(sw)\right\} \qquad w \in [-\pi, \pi]$$

where $R(s)$ denotes the sample autocovariance function, i.e.

$$R(s) = \frac{1}{n}\sum_{i=1}^{n-s}(X_i - \bar{X})(X_{i+s} - \bar{X}) \qquad s = 0, \pm 1, \ldots, \pm(n-1)$$

and $\bar{X}$ is the sample mean. For negative values of d and with $\{\varepsilon_t\}$ a purely random process with $E(\varepsilon_t^2)$ finite, we find that the well-known relationship between the periodograms of $\{X_t\}$ and $\{\varepsilon_t\}$ can be applied to the ARIMA (p, d, q) process when $d < 0$. This results in the following lemma (cf. Hassler, 1993).

LEMMA 1.1. *Let* $X_t = (1 - B)^{-d}\{\Theta(B)/\Phi(B)\}\varepsilon_t$ *be an* $ARIMA(p, d, q)$ *process with* $d < 0$ *and let* $E(\varepsilon_t^4)$ *be finite. Then the periodogram of* $\{X_t\}$, $I_x(w)$, *is asymptotically given by*

$$I_x(w_j) \approx 2\pi\frac{f(w)}{\sigma_\varepsilon^2} I_\varepsilon(w_j),$$

$I_\varepsilon(w)$ *being the periodogram of* $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$, *and the* $\{I_x(w_j)\}$, $w_j \in [0, \pi]$, *are asymptotically independent with distributions given by*

$$I_x(w_j) = \begin{cases} \frac{1}{2}f(w_j)\chi_2^2 & j \neq 0, [n/2] \\ f(w_j)\chi_1^2 & j = 0, [n/2]. \end{cases}$$

(It should be noted that the results of Kunsch (1986) indicate that these asymptotic distributions may not hold for frequencies very close to zero. We will, nevertheless, use the above results as an approximation *to the distribution of the periodogram ordinates throughout the whole frequency range.)*

Using Lemma 1.1 we now obtain the following lemma.

LEMMA 1.2. *Let* $\{X_t\}$ *be an* $ARIMA(p, d, q)$ *process of the form (1) with* $d < 0$ *and* ε_t *a white noise process. When* $n \to \infty$ *the sequence* $\{-\ln[I_x(w_j)/f(w_j)]\}$, $w_j = 2\pi j/n$, $j = 1, 2, \ldots, [n/2] - 1$, *follows independent Gumbel distributions with mean 0.577216 (Euler's constant) and variance* $\pi^2/6$.

PROOF. For $d < 0$, using Lemma 1.1 we find that, as $n \to \infty$, $Y = 2I_x(w_j)/f(w_j)$, $w_j \neq 0$, π, has a χ^2 distribution with two degrees of freedom. Thus, $Y \approx e^{1/2}$. Let $F_z(z, \alpha, \beta) = \exp(-e^{-(z-\alpha)/\beta})$ be the cumulative distribution function of the Gumbel distribution. If Z is a random variable having this distribution, then $E(Z) = \alpha + \beta\delta$ (δ is Euler's constant) and $\mathrm{var}(Z) = \pi^2\beta/6$ (Mood *et al.*, 1974). Now let F_x be the cumulative distribution function of $\{-\ln[I_x(w_j)/f(w_j)]\}$. Then $F_x = P[-\ln\{I_x(w_j)/f(w_j)\} \leqslant x] = P[-\ln\{2I_x(w_j)/f(w_j)\} \leqslant x - \ln 2] = P(Y > e^{-x+\ln 2}) = \exp(-e^{-x})$. Thus, F_x is a cumulative distribution of the same form as F_z, with $\alpha = 0.0$ and $\beta = 1.0$. The independence is also guaranteed by Lemma 1.1.

The estimator $\hat{d}_{\mathrm{p}}$

Now we return to equation (3). Adding $\ln\{I(w_j)\}$ to both sides we have

$$\ln\{I(w_j)\} = \ln\{f_u(0)\} - d\ln\left\{2\sin\left(\frac{w_j}{2}\right)\right\}^2 + \ln\left\{\frac{f_u(w_j)}{f_u(0)}\right\} + \ln\left\{\frac{I(w_j)}{f(w_j)}\right\}. \tag{4}$$

Now if the upper limit of j, say $g(n)$, is chosen so that $g(n)/n \to 0$ as $n \to \infty$ and if w_j is 'near zero', say $w_j \leq w_{g(n)}$ where $w_{g(n)}$ is 'small', then the term $\ln\{f_u(w_j)/f_u(0)\}$ is negligible compared with the others on the right-hand side, so we can re-write (4) as

$$\ln\{I(w_j)\} \approx \ln\{f_u(0)\} - d\ln\left\{2\sin\left(\frac{w_j}{2}\right)\right\}^2 + \ln\left\{\frac{I(w_j)}{f(w_j)}\right\}. \tag{5}$$

This now has the form of a simple regression equation, namely

$$y_j = a + bx_j + e_j \qquad j = 1, 2, \ldots, g(n) \tag{6}$$

where $y_j = \ln\{I(w_j)\}$, $x_j = \ln\{2\sin(w_j/2)\}^2$, $e_j = \ln\{I(w_j)/f(w_j)\} + c$, $b = -d$, $a = \ln\{f_u(0)\} - c$ and $c = E[-\ln\{I(w_j)/f(w_j)\}]$.

From Lemma 1.2, when $d \in (-0.5, 0.0)$ the members of the sequence $\{\ln[I(w_j)/f(w_j)]\}$, $j = 1, 2, \ldots, g(n)$, are approximately independent Gumbel random variables with mean -0.577216 and variance $\pi^2/6$. Thus, the members of the sequence $\{e_j\}$ are approximately independent Gumbel random variables with mean 0 and variance $\pi^2/6$. This suggests estimating d by least-squares regression of $y_1, y_2, \ldots, y_{g(n)}$ on $x_1, x_2, \ldots, x_{g(n)}$ where $g(n)$ is a function of n chosen as described above.

This leads to the estimator

$$\hat{b} = \frac{\sum_{i=1}^{g(n)}(x_i - \bar{x})y_i}{\sum_{i=1}^{g(n)}(x_i - \bar{x})^2}.$$

The estimator of d using the periodogram in the regression method is then given by $\hat{d}_{\mathrm{p}} = -\hat{b}$ and we have $E(\hat{d}_{\mathrm{p}}) = d$ and

$$\operatorname{var}(\hat{d}_{\mathrm{p}}) = \frac{\pi^2}{6\sum_{i=1}^{g(n)}(x_i - \bar{x})^2}.$$

The asymptotic distribution of $\hat{d}_{\mathrm{p}}$ is given by Geweke and Porter-Hudak (1983). They also suggest taking $g(n) = n^{\alpha}$, $0 < \alpha < 1$, and that if $\lim\{(\ln n)^2/g(n)\} = 0$ then $(\hat{d}_{\mathrm{p}} - d)/\{\operatorname{var}(\hat{d}_{\mathrm{p}})\}^{1/2}$ is asymptotically distributed as $\mathrm{N}(0, 1)$.

3.2. *The estimator of d using the smoothed periodogram*

We now consider the use of a consistent estimate of the spectral density, i.e. a smoothed periodogram using the Parzen lag window, for estimating the parameter d. Let $f_s(w)$ denote a smoothed periodogram of the form

$$f_s(w_j) = \frac{1}{2\pi} \sum_{s=-m}^{m} k\left(\frac{s}{m}\right) R(s) \cos(sw_j) \tag{7}$$

where $k(u)$ is called the lag window generator, a fixed continuous even function in the range $-1 < u < 1$, with $k(0) = 1$ and $k(-u) = k(u)$. The parameter m (usually referred to as the 'truncation point') is a function of n (sample size) chosen such that, as $n \to \infty$, $m \to \infty$ and $(m/n) \to 0$. (We could choose, for example, $m = n^{\beta}$, $0 < \beta < 1$). The Parzen lag window generator has the following form:

$$k(u) = \begin{cases} 1 - 6u^2 + 6|u|^3 & |u| \leqslant 1/2 \\ 2(1 - |u|)^3 & -1/2 < u \leqslant 1 \\ 0 & |u| > 1. \end{cases} \tag{8}$$

(We choose the Parzen lag window because it has the property that it always produces positive estimates of the spectral density.) By Priestley (1981) we have that $f_s(w_j)$ is asymptotically unbiased with variance given by

$$\operatorname{var}\{f_s(w_j)\} \approx \begin{cases} 0.539\,285\left(\dfrac{m}{n}\right) f^2(w_j) & w_j \neq 0, \pi \\ 1.078\,56\left(\dfrac{m}{n}\right) f^2(w_j) & w_j = 0, \pi \end{cases} \tag{9}$$

and $\lim_{n\to\infty} (n/m) \operatorname{cov}\{f_s(w_1), f_s(w_2)\} = 0$, $w_1 \neq \pm w_2$.

According to Anderson ((1971), ch. 9), under general conditions it can be shown that the smoothed spectral estimate $f_s(w_j)$ and $\ln\{f_s(w_j)\}$ each have an asymptotic normal distribution. Using the asymptotic results for $f_s(w)$ we arrive at the following lemma.

LEMMA 1.3. *Let $f(w)$ be the spectral density function of the ARIMA(p, d, q) model ($d \in (-0.5, 0.0)$) and $f_s(w)$ its estimator of the form (7), $k(s/m)$ being the Parzen lag window generator. Then,* $\ln\{f_s(w)/f(w)\}$ *has an asymptotically normal distribution with mean 0 and variance given by*

$$\operatorname{var}\left[\ln\left\{\frac{f_s(w_j)}{f(w_j)}\right\}\right] \approx \begin{cases} 0.539\,285\left(\dfrac{m}{n}\right) & w \neq 0, \pi \\ 1.078\,56\left(\dfrac{m}{n}\right) & w = 0, \pi. \end{cases} \tag{10}$$

Returning to expression (3) and taking into account the asymptotic results for $\ln\{f_s(w)/f(w)\}$ which we have just described, we can still write the regression equation based on the smoothed periodogram in the form

$$\ln\{f_s(w_j)\} = \ln\{f_u(0)\} - d\ln\left\{2\sin\left(\frac{w_j}{2}\right)\right\}^2 + \ln\left\{\frac{f_s(w_j)}{f(w_j)}\right\} + \ln\left\{\frac{f_u(w_j)}{f_u(0)}\right\}. \tag{11}$$

Restricting the range of j to $1 \leqslant j \leqslant g(n)$, and choosing $g(n)$ as previously, we can re-write (11) in the form

$$\ln\{f_s(w_j)\} \approx \ln\{f_u(0)\} - d\ln\left\{2\sin\left(\frac{w_j}{2}\right)\right\}^2 + \ln\left\{\frac{f_s(w_j)}{f(w)}\right\}. \tag{12}$$

Equation (12) retains the same simple linear regression form as before, i.e. it takes the form $y_j = a + bx_j + e_j$, $j = 1, 2, 3, \ldots, g(n)$, where now $y_j = \ln\{f_s(w_j)\}$, $b = -d$, $x_j = \ln\{2\sin(w_j/2)\}^2$, $e_j = \ln\{f_s(w_j)/f(w_j)\}$ and $a = \ln\{f_u(0)\}$. As previously noted, when $-0.5 < d < 0.0$ in the ARIMA(p, d, q) process the e_j are asymptotically uncorrelated with mean 0 and variance given by (10). This suggests that the regression method for estimating d may be applied to equation (3) where $g(n)$ is chosen as before.

The estimator $\hat{d}_{sp}$

The estimator of d obtained by the regression method using the smoothed periodogram with the Parzen lag window is given by $\hat{d}_{sp} = -\hat{b}$, where

$$\hat{b} = \frac{\sum_{i=1}^{g(n)}(x_i - \bar{x})y_i}{\sum_{i=1}^{g(n)}(x_i - \bar{x})^2}$$

and the corresponding variance is

$$\operatorname{var}(\hat{d}_{sp}) \approx 0.539\,28\frac{m}{n\sum_{i=1}^{g(n)}(x_i - \bar{x})^2} \qquad w \neq 0, \pi.$$

Note that these results can be generalized for any spectral estimate using a lag window generator which satisfies the conditions of Theorem 9.2.1 in Anderson (1971). The variance of e_j will then be approximately $(m/n)\int k^2(u)\,\mathrm{d}u$.

3.3. *The Hurst coefficient method*

This method is based on the studies of McLeod and Hipel (1978) and O'Conell (1974) which are related to the original studies of Hurst (1951). We will summarize the method here but more details can be found in the works cited above. Let $\{X_t\}$, $t = 1, 2, \ldots, n$, denote the observations and let $\bar{X}$ denote the sample mean. Define the kth general partial sum as $S'_k = S'_{k-1} + (X_k - \alpha\bar{X})$ where $S'_0 = 0$ and α is a constant satisfying $0 \leqslant \alpha \leqslant 1$.

When $\alpha = 1$ the sum S'_k is given by $S^*_k = S^*_{k-1} + (X_k - \bar{X}) = \sum_{i=1}^{k} X_i - k\bar{X}$, $k = 1, 2, 3, \ldots, n$, where $S^*_0 = 0$ and $S^*_n = 0$. The *adjusted range* R^*

is defined as $R_n^* = M_n^* - m_n^*$ where $M_n^* = \max(0, S_1^*, \ldots, S_n^*)$ and $m_n^* = \min(0, S_1^*, \ldots, S_n^*)$. M_n^* and m_n^* are known as the *adjusted surplus* and the *adjusted deficit* respectively. *The rescaled adjusted range* (RAR) is defined as

$$R_n = \frac{R_n^*}{D_n^*}$$

where

$$D_n^* = \left\{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}\right\}^{1/2}$$

is the sample standard deviation.

McLeod and Hipel (1978) found by simulation studies that for indendently distributed random variables the RAR does not significantly depend on the underlying distribution of the random variables but is a function only of the sample size. Interest in the statistic R_n was stimulated by Hurst's studies (1951, 1956) of the long-term storage requirements on the River Nile. For approximately 900 annual time series comprising streamflow and precipitation records, stream and lake levels, tree rings, mud waves, atmospheric pressure and sunspots, Hurst found R_n to vary with n according to $R_n \sim n^h$ where h is a constant called the *generalized Hurst coefficient*. The expression of R_n can thus be written in general form $R_n = bn^h$ where b is a coefficient that is not a function of n. Hurst estimated the coefficient h by the coefficient k from the following relationship (taking b in the above equation to have a value $(1/2)^{\mathrm{h}}$):

$$R_n = \left(\frac{n}{2}\right)^k. \tag{13}$$

By taking logarithms of (13), k, the estimate of h, is evaluated for each time series as

$$k = \frac{\ln R_n^* - \ln D_n^*}{\ln(n/2)}. \tag{14}$$

Hurst (1951) gave a theoretical value for the expected value of R_n^* for an independent normal process, which is $E(R_n^*) = (n\pi/2)^{0.5}\sigma$ where σ^2 is the variance of the process. (Note that the expression (13) which was used by Hurst for estimating h was apparently suggested by the equation for $E(R_n^*)$ where the term $n/2$ appears.) Without invoking the assumption of normality for the underlying process, Feller (1951) derived an asymptotic result of $E(R_n^*)$ using the theory of Brownian motion (see, for example Hosking, 1982, and Geweke and Porter-Hudak, 1983). Other estimates of the parameter h have been formulated and are summarized in McLeod and Hipel (1978). We will concentrate in our studies on estimating h by the parameter k using the form (14).

The fractional Gaussian noise (FGN) process, which was first studied by Mandelbrot (1965), may be used as a theoretical basis for Hurst's law. Mandelbrot and Wallis (1969a, b, c) and Mandelbrot and Vann Ness (1968) proposed a rigorous theory (which will not be presented here) of FGN which adequately accounted for the Hurst phenomenon (a summary can be found in O'Conell, 1974).

The FGN process has a parameter H (as employed by those authors) which must lie in $(0, 1)$ and their research shows that if $1/2 < H < 1$ then the FGN process has long memory properties. Also the authors claimed that for the FGN process the statistic RAR varies with n as $R_n^*/D_n^* \sim n^H$. Hence, the parameter H in FGN is often estimated by the Hurst coefficient estimate k (cf. result (14)). A summary of FGN is found in Hosking (1982), McLeod and Hipel (1978) and O'Conell (1974).

The ARIMA(p, d, q) process is tightly connected to the FGN process. Hosking (1982, 1984) shows that the ARIMA$(0, d, 0)$ model and the FGN process with parameter H have similar correlation structures. For large k, ρ_k behaves like k^{2H-2} for the FGN process and like k^{2d-1} for the ARIMA$(0, d, 0)$ model but the constant of proportionality is different for the two processes. The relationship between the two processes, FGN and ARIMA$(0, d, 0)$, is further developed by Geweke and Porter-Hudak (1983). They derive the spectral density function for a general fractional Gaussian noise (GFGN) with parameter $H \in (0, 1)$ and prove that it is the same as that for the ARIMA(p, d, q) $(d \in (-0.5, 0.5))$ process where the parameter $d = H - 0.5$. On the basis of Geweke and Porter-Hudak's result the Hurst coefficient method for estimating d in the ARIMA(p, d, q) model leads to the estimate

$$\hat{d}_{\mathrm{h}} = \hat{H} - 0.5 \tag{15}$$

where $\hat{H}$ is computed using expression (14).

4. EXPERIMENTAL RESULTS COMPARING THE ESTIMATORS $\hat{d}_{\mathrm{sp}}$, $\hat{d}_{\mathrm{p}}$ AND $\hat{d}_{\mathrm{h}}$

Three estimators of d $\hat{d}_{\mathrm{sp}}$, $\hat{d}_{\mathrm{p}}$ and $\hat{d}_{\mathrm{h}}$ will be compared for different cases of the ARIMA(p, d, q) process, i.e. with different values of $d \in (-0.5, 0.5)$, p and q in the model. Also the effect of the sample size n will be examined.

In the regression equation, using both the periodogram and the smoothed periodogram, different forms of the function $g(n) = n^{\alpha}$, $0 < \alpha < 1$, will be considered in order to examine the effect of α in the regression equation estimate of d. The truncation point $m (m = n^{\beta}$, $0 < \beta < 1)$ will also be discussed in order to see its effect on the values of $\hat{d}_{\mathrm{sp}}$.

Our experimental results will be shown in tables using the following notation:

(i) $\bar{d}_i - d$, $i \equiv$ sp, p and h, refers to the mean bias of the estimators $\hat{d}_{\mathrm{sp}}$, $\hat{d}_{\mathrm{p}}$ and $\hat{d}_{\mathrm{h}}$ respectively.

(ii) $\widehat{\mathrm{MSE}}_i$ is the estimated mean square error (MSE) of $\hat{d}_i$.

(iii) $\widehat{\mathrm{var}}_i$ is the estimated variance of $\hat{d}_i$.

Samples of the ARIMA(p, d, q) process were generated from the approximation method suggested by Hosking (1981), taking the white noise process $\{\varepsilon_t\}$ to have an N(0, 1) distribution. We have taken a sample size of 300.

In Tables I and II the value of α is fixed to be 0.5 (so that $g(n) = n^{\alpha} = 17$) in the regression equation, and the truncation point in the Parzen lag window is chosen as $m = n^{\beta} = 169$ (i.e. $\beta = 0.9$). The choice of the values of $g(n)$ and m will be discussed with the results in Tables IV and V respectively. The value of d lies in the interval $(-0.5, 0.5)$ and these values are described in the tables as well as the values of the parameters of the ARIMA(p, q) model. The results shown in this group refer to the mean value over 30 replications of the biases for $\hat{d}_{\mathrm{sp}}$, $\hat{d}_{\mathrm{p}}$ and $\hat{d}_{\mathrm{h}}$ and the estimated MSE for each estimator.

In Table I, the model considered has an AR(1) component and the bias of $\hat{d}_{\mathrm{sp}}$ is smaller than those of $\hat{d}_{\mathrm{p}}$ and $\hat{d}_{\mathrm{h}}$ is almost all cases. The bias of $\hat{d}_{\mathrm{sp}}$ is

TABLE I
ARIMA$(1, d, 0)$, $(1-\phi B)(1-B)^d X_t = \varepsilon_t$

	$d = -0.2$		$d = 0.1$		$d = 0.3$		$d = 0.4$	
	$d_i - d$	$\widehat{\mathrm{MSE}}_i$	$d_i - d$	$\widehat{\mathrm{MSE}}_i$	$d_i - d$	$\widehat{\mathrm{MSE}}_i$	$d_i - d$	$\widehat{\mathrm{MSE}}_i$
$\phi = -0.4$								
$i =$ sp	0.027	0.018	0.006	0.018	−0.001	0.019	−0.004	0.021
$i =$ p	0.045	0.036	0.033	0.032	0.028	0.030	0.026	0.030
$i =$ h	0.128	0.017	−0.014	0.002	−0.095	0.011	−0.139	0.020
$\phi = -0.2$								
$i =$ sp	0.029	0.018	0.008	0.018	0.001	0.019	−0.002	0.020
$i =$ p	0.045	0.037	0.036	0.033	0.030	0.031	0.028	0.030
$i =$ h	0.168	0.029	0.024	0.002	−0.065	0.006	−0.114	0.020
$\phi = 0.0$								
$i =$ sp	0.032	0.019	0.011	0.019	0.004	0.019	0.001	0.020
$i =$ p	0.049	0.038	0.040	0.035	0.034	0.032	0.032	0.030
$i =$ h	0.206	0.043	0.058	0.005	−0.039	0.0035	−0.095	0.010
$\phi = 0.2$								
$i =$ sp	0.039	0.019	0.018	0.019	0.011	0.020	0.008	0.020
$i =$ p	0.056	0.039	0.047	0.036	0.041	0.033	0.039	0.030
$i =$ h	0.025	0.061	0.091	0.010	−0.016	0.002	−0.077	0.010
$\phi = 0.4$								
$i =$ sp	0.054	0.021	0.034	0.020	0.027	0.021	0.025	0.020
$i =$ p	0.072	0.041	0.062	0.038	0.057	0.034	0.056	0.030
$i =$ h	0.289	0.085	0.126	0.017	0.007	0.002	−0.059	0.010
$\phi = 0.6$								
$i =$ sp	0.098	0.028	0.079	0.026	0.073	0.025	0.07	0.030
$i =$ p	0.116	0.049	0.105	0.047	0.101	0.040	0.10	0.040
$i =$ h	0.340	0.119	0.166	0.029	0.034	0.003	−0.039	0.004
$\phi = 0.7$								
$i =$ sp	0.151	0.042	0.133	0.037	0.127	0.036	0.124	0.040
$i =$ p	0.169	0.064	0.156	0.058	0.156	0.053	0.150	0.050
$i =$ h	0.378	0.145	0.192	0.038	0.049	0.004	−0.028	0.003

TABLE II
ARIMA(0, d, 1), $(1-B)^d X_t = (1-\theta B)\varepsilon_t$

	$d=-0.2$		$d=0.1$		$d=0.3$		$d=0.4$	
	d_i-d	$\widehat{\text{MSE}}_i$	d_i-d	$\widehat{\text{MSE}}_i$	d_i-d	$\widehat{\text{MSE}}_i$	d_i-d	$\widehat{\text{MSE}}_i$
$\theta=-0.4$								
$i=$ sp	0.037	0.019	0.016	0.019	0.009	0.020	0.007	0.021
$i=$ p	0.053	0.039	0.045	0.036	0.040	0.034	0.037	0.031
$i=$ h	0.260	0.069	0.100	0.011	−0.010	0.002	−0.070	0007
$\theta=-0.2$								
$i=$ sp	0.035	0.019	0.015	0.019	0.008	0.019	0.005	0.021
$i=$ p	0.052	0.039	0.044	0.036	0.038	0.033	0.036	0.031
$i=$ h	0.239	0.058	0.084	0.009	−0.021	0.002	−0.081	0.009
$\theta=0.2$								
$i=$ sp	0.025	0.018	0.004	0.018	−0.003	0.019	−0.006	0.021
$i=$ p	0.041	0.037	0.032	0.033	0.027	0.030	0.025	0.029
$i=$ h	0.163	0.028	0.017	0.002	−0.070	0.007	−0.119	0.017
$\theta=0.4$								
$i=$ sp	0.005	0.020	−0.012	0.018	−0.019	0.019	−0.022	0.021
$i=$ p	0.024	0.036	0.015	0.031	0.010	0.029	0.008	0.028
$i=$ h	0.110	0.013	−0.039	0.003	−0.121	0.017	−0.163	0.029
$\theta=0.6$								
$i=$ sp	−0.016	0.018	−0.058	0.021	−0.066	0.022	−0.069	0.025
$i=$ p	−0.023	0.034	−0.031	0.034	−0.035	0.028	−0.037	0.028
$i=$ h	0.054	0.004	−0.110	0.014	−0.197	0.040	−0.240	0.058
$\theta=0.7$								
$i=$ sp	−0.057	0.022	−0.110	0.031	−0.120	0.032	−0.124	0.035
$i=$ p	−0.078	0.037	−0.085	0.036	−0.089	0.035	−0.091	0.034
$i=$ h	0.026	0.001	−0.159	0.026	−0.250	0.064	−0.290	0.085

about half that of $\hat{d}_p$. However, the estimator $\hat{d}_h$ shows a smaller bias than that of $\hat{d}_{sp}$ in only five cases, when d increases ($d=0.3$, 0.4) and the coefficient ϕ also increases ($\phi=0.4$, 0.6, 0.7). It can be seen that as the values of the coefficient ϕ increases positively and become closer to non-stationary conditions (i.e. close to 1), the biases in both $\hat{d}_{sp}$ and $\hat{d}_p$ also increase. The bias of $\hat{d}_h$ is largest when $d=-0.2$ and $d=0.1$; when the process is 'short memory' ($d=-0.2$) it can be seen that $\hat{d}_h$ is very sensitive, resulting in a higher bias, but when the process becomes 'long memory' and especially when the coefficient ϕ also becomes larger, the bias decreases.

The $\widehat{\text{MSE}}$ of $\hat{d}_{sp}$ is always smaller than that of $\hat{d}_p$ (it tends to be approximately half as large): their values do not vary too much, lying in the ranges (0.018, 0.042) and (0.03, 0.064) respectively. However, the $\widehat{\text{MSE}}$ of $\hat{d}_h$ sometimes shows smaller values than that of $\hat{d}_{sp}$.

When an MA(1) component is included (Table II) the bias of $\hat{d}_{sp}$ is again usually smaller than those of $\hat{d}_p$ and $\hat{d}_h$. The estimator $\hat{d}_p$ shows smaller bias than in the other six cases ($d=0.3$, 0.4 and $\theta=0.4$, 0.6, 0.7). When $|\theta|$ gets closer to 1 (i.e. closer to non-stationary conditions) the biases of $\hat{d}_{sp}$ and $\hat{d}_p$

TABLE III

ARIMA(1, d, 0), $(1-\phi B)(1-B)^d X_t = \varepsilon_t$, $d = 0.3$

	n = 200		n = 300		n = 400	
	$d_i - d$	$\widehat{\text{MSE}}_i$	$d_i - d$	$\widehat{\text{MSE}}_i$	$d_i - d$	$\widehat{\text{MSE}}_i$
$\phi = -0.4$						
i = sp	−0.042	0.031	−0.001	0.019	−0.005	0.017
i = p	0.039	0.033	0.003	0.030	0.021	0.028
i = h	−0.109	0.015	−0.095	0.011	−0.104	0.012
$\phi = -0.2$						
i = sp	−0.039	0.031	0.001	0.019	−0.003	0.017
i = p	0.043	0.034	0.030	0.031	0.022	0.029
i = h	−0.077	0.009	−0.064	0.006	−0.075	0.007
$\phi = 0.0$						
i = sp	−0.033	0.031	0.004	0.019	−0.001	0.016
i = p	0.047	0.034	0.034	0.032	0.025	0.029
i = h	−0.049	0.005	−0.039	0.004	−0.051	0.004
$\phi = 0.2$						
i = sp	−0.022	0.029	0.011	0.020	0.005	0.016
i = p	0.056	0.034	0.041	0.033	0.031	0.031
i = h	−0.024	0.003	−0.016	0.002	−0.029	0.003
$\phi = 0.4$						
i = sp	0.003	0.029	0.027	0.021	0.017	0.017
i = p	0.082	0.038	0.057	0.035	0.043	0.033
i = h	0.003	0.002	0.007	0.002	−0.006	0.002
$\phi = 0.6$						
i = sp	0.068	0.034	0.073	0.025	0.053	0.020
i = p	0.150	0.052	0.101	0.040	0.079	0.040
i = h	0.033	0.003	0.034	0.003	0.018	0.002

increase. In only one instance ($d = -0.2$, $\theta = 0.7$) does $\hat{d}_h$ show a smaller bias than those of $\hat{d}_{sp}$ and $\hat{d}_p$. It appears that when the coefficient θ is negative the biases of $\hat{d}_{sp}$ and $\hat{d}_p$ become smaller as the value of d increases, and the bias of $\hat{d}_h$ becomes irregular. As the value of the coefficient θ increases for $d > 0$, the biases of all three estimators become larger. Analysing the values of $\widehat{\text{MSE}}$, it again becomes apparent that the $\widehat{\text{MSE}}$ of $\hat{d}_{sp}$ is always smaller than that of $\hat{d}_p$, by almost a half. However, the $\widehat{\text{MSE}}$ of d_h shows a smaller value than that of d_{sp} in only some instances. When the coefficient θ has a negative value, the $\widehat{\text{MSE}}$ of $\hat{d}_h$ is irregular for all values of d, yet as θ becomes positive the $\widehat{\text{MSE}}$ increases as d increases.

It has already been noted that as the coefficients ϕ and θ increase the biases of the estimators $\hat{d}_{sp}$ and $\hat{d}_p$ increase substantially. This may be explained by the fact that the term $\ln\{f_u(w)/f_u(0)\}$, which involves the unknown autoregressive and moving-average parameters, in equations (4) and (11) has been dropped.

In Table III, where different sample sizes (200, 300 and 400) are applied, it can be seen that in almost all cases the biases of both $\hat{d}_{sp}$ and $\hat{d}_p$ decrease as

TABLE IV
ARIMA(1, d, 0), $d = 0.3$, $n = 300$, $g(n) = n^\alpha$, $m = 169$

	$\alpha = 0.5$ $(g(n) = 17)$			$\alpha = 0.7$ $(g(n) = 54)$			$\alpha = 0.8$ $(g(n) = 95)$		
	$d - d_i$	$\widehat{\text{var}}_i$	$\widehat{\text{MSE}}_i$	$d - d_i$	$\widehat{\text{var}}_i$	$\widehat{\text{MSE}}_i$	$d - d_i$	$\widehat{\text{var}}_i$	$\widehat{\text{MSE}}_i$
$\phi = -0.4$									
$i = \text{sp}$	−0.001	0.019	0.019	−0.044	0.0046	0.0065	−0.12	0.003	0.017
$i = \text{p}$	0.028	0.029	0.030	−0.032	0.0067	0.0077	−0.11	0.005	0.017
$i = \text{h}$	−0.095	0.002	0.011						
$\phi = -0.2$									
$i = \text{sp}$	0.001	0.019	0.019	−0.032	0.0046	0.006	−0.075	0.0032	0.0088
$i = \text{p}$	0.030	0.029	0.031	−0.019	0.0068	0.007	−0.067	0.0055	0.0098
$i = \text{h}$	−0.065	0.002	0.006						
$\phi = 0.0$									
$i = \text{sp}$	0.004	0.019	0.019	−0.007	0.0046	0.0046	−0.009	0.003	0.0031
$i = \text{p}$	0.034	0.030	0.032	0.004	0.0070	0.0070	−0.001	0.005	0.0054
$i = \text{h}$	−0.039	0.002	0.0035						
$\phi = 0.2$									
$i = \text{sp}$	0.011	0.020	0.020	0.038	0.0046	0.0060	0.086	0.003	0.010
$i = \text{p}$	0.041	0.031	0.033	0.047	0.0070	0.0092	0.093	0.005	0.0141
$i = \text{h}$	−0.016	0.002	0.002						
$\phi = 0.4$									
$i = \text{sp}$	0.027	0.020	0.021	0.117	0.0045	0.018	0.216	0.003	0.050
$i = \text{p}$	0.057	0.031	0.034	0.128	0.0070	0.023	0.223	0.005	0.055
$i = \text{h}$	0.007	0.002	0.002						
$\phi = 0.6$									
$i = \text{sp}$	0.073	0.020	0.025	0.26	0.0044	0.072	0.39	0.0036	0.156
$i = \text{p}$	0.101	0.030	0.040	0.27	0.0074	0.080	0.40	0.0055	0.162
$i = \text{h}$	0.034	0.019	0.003						
$\phi = 0.7$									
$i = \text{sp}$	0.128	0.020	0.036	0.37	0.0045	0.14	0.49	0.0031	0.24
$i = \text{p}$	0.156	0.028	0.053	0.38	0.0072	0.15	0.50	0.0058	0.26
$i = \text{h}$	0.049	0.002	0.004						
	$\text{var}(\hat{d}_{sp}) = 0.0075$			$\text{var}(\hat{d}_{sp}) = 0.0018$			$\text{var}(\hat{d}_{sp}) = 0.0011$		
	$\text{var}(\hat{d}_{p}) = 0.041$			$\text{var}(\hat{d}_{p}) = 0.0099$			$\text{var}(\hat{d}_{p}) = 0.0057$		

the sample size increases. Again $\hat{d}_{sp}$ has a smaller bias than $\hat{d}_p$. The $\widehat{\text{MSE}}$s of both $\hat{d}_{sp}$ and $\hat{d}_p$ are reduced as the sample size increases. The $\widehat{\text{MSE}}$ of $\hat{d}_{sp}$ is reduced by approximately a half as n increases from 200 to 400. The estimator $\hat{d}_h$ does not have a regular behavioural pattern as sometimes its bias increases as the sample size increases. Its $\widehat{\text{MSE}}$ value does not vary too much. Again, in these results we take $g(n) = n^{0.5}$ and $m = n^{0.9}$.

In Table IV different values of α (0.5, 0.7, 0.9) are applied in the regression equation, which produces a regression based on 17, 54 and 95 observations. The results suggest that as the number of observations of the regression equation increases the estimated variances of both $\hat{d}_{sp}$ and $\hat{d}_p$ decrease, while their biases increase. The theoretical variances of $\hat{d}_{sp}$ and $\hat{d}_p$ are given at the bottom of the table for each value of α. The variance estimate of $\hat{d}_{\mathrm{p}}$ is always larger than that of $\hat{d}_{\mathrm{sp}}$. However, in some cases the bias of $\hat{d}_{\mathrm{p}}$ is smaller than that of $\hat{d}_{\mathrm{sp}}$.

In Table V, different values of the truncation point m in the Parzen lag window are used. These values are 30, 54 and 169. The results show that the

TABLE V

ARIMA(1, d, 0), $d = 0.3$, $m = n^{\beta}$, $n = 300$, $g(n) = 17$

	$\beta = 0.6$ ($m = 30$)			$\beta = 0.7$ ($m = 54$)			$\beta = 0.9$ ($m = 169$)		
	$d - d_i$	$\widehat{\text{var}}_i$	$\widehat{\text{MSE}}_i$	$d - d_i$	$\widehat{\text{var}}_i$	$\widehat{\text{MSE}}_i$	$d - d_i$	$\widehat{\text{var}}_i$	$\widehat{\text{MSE}}_i$
$\phi = -0.5$									
i = sp	−0.029	0.012	0.013	−0.029	0.018	0.019	−0.0019	0.019	0.019
i = p	0.027	0.029	0.030						
i = h	−0.114	0.002	0.015						
$\phi = -0.4$									
i = sp	−0.024	0.013	0.014	−0.027	0.019	0.020	−0.001	0.019	0.019
i = p	0.028	0.029	0.030						
i = h	−0.095	0.002	0.011						
$\phi = -0.2$									
i = sp	−0.027	0.015	0.016	−0.024	0.020	0.020	0.001	0.019	0.019
i = p	0.030	0.029	0.031						
i = h	−0.065	0.002	0.006						
$\phi = 0.0$									
i = sp	−0.018	0.015	0.015	−0.019	0.019	0.019	0.004	0.019	0.019
i = p	0.034	0.030	0.032						
i = h	−0.039	0.002	0.0035						
$\phi = 0.2$									
i = sp	−0.014	0.014	0.014	−0.013	0.019	0.019	0.011	0.020	0.020
i = p	0.041	0.031	0.033						
i = h	−0.016	0.002	0.0022						
$\phi = 0.4$									
i = sp	0.0023	0.015	0.015	0.0044	0.018	0.018	0.027	0.020	0.021
i = p	0.0570	0.031	0.034						
i = h	0.0070	0.002	0.002						
$\phi = 0.5$									
i = sp	0.020	0.015	0.015	0.020	0.018	0.018	0.044	0.020	0.022
i = p	0.073	0.031	0.036						
i = h	0.020	0.0019	0.002						
$\phi = 0.6$									
i = sp	0.036	0.015	0.016	0.054	0.020	0.022	0.073	0.020	0.025
i = p	0.101	0.030	0.040						
i = h	0.034	0.002	0.003						
$\phi = 0.7$									
i = sp	0.083	0.016	0.022	0.096	0.018	0.027	0.127	0.020	0.036
i = p	0.156	0.029	0.053						
i = h	0.049	0.002	0.004						
	$\text{var}(\hat{d}_{\mathrm{sp}}) = 0.0014$			$\text{var}(\hat{d}_{\mathrm{sp}}) = 0.0024$			$\text{var}(\hat{d}_{\mathrm{sp}}) = 0.0075$		
	$\text{var}(\hat{d}_{\mathrm{p}}) = 0.041$								

estimated variance increases as m increases. However, the largest values of $\widehat{\mathrm{var}}(\hat{d}_{sp})$ are always smaller than the estimated variance of $\hat{d}_p$, which remains constant as m varies. Also, the results suggest that the bias of $\hat{d}_{sp}$ is reduced in almost all cases as m increases; when ϕ becomes positive and larger the bias increases. These results reflect the well-known fact that the bias of the spectral estimator decreases with m, while the variance increases (see, for example, Priestley, 1981).

In Table VI we present results of testing the null hypothesis H_0: $d = 0.0$ while the samples are generated for values of d (which are written in the table) different from zero. In this situation the results obtained show the percentage of times H_0 was rejected when it is false. As we have seen in a previous section, $\hat{d}_{sp} - d$ and $\hat{d}_p - d$ have asymptotic normal distributions. Again, the asymptotic theory described for these estimates is valid only for $d < 0$, but we also consider the case when $d > 0$ in our experimental studies. The results are shown in Table VI using a test based on the asymptotic normal distributions with a 5% significance level. The samples are from the ARIMA$(1, d, 0)$ model and $\varepsilon_t \sim N(0, 1)$. The sample size is 300 and 100 replications of each experiment. The number of observations in the regression is equal to 17 ($n^{0.5}$) and the truncation point m in the Parzen lag window (for the smoothed periodogram function) is $n^{0.9} = 169$. The results show that a high probability of rejecting H_0 is generally obtained when estimator $\hat{d}_{sp}$ is used. It is clear that the highest percentile of rejecting H_0 occurs when $\hat{d}_{sp}$ is applied. Also, as expected, the percentile becomes larger for both estimators as d and ϕ increase. These computer experiments show that the smoothed periodogram regression may be superior to the periodogram regression when we wish to discriminate between an ARIMA$(p, 0, q)$ model and an ARIMA(p, d, q) model with $d \neq 0$.

Note that after submitting this paper it came to my attention that a similar approach to the estimation of the fractional difference parameter based on the regression analysis of the smoothed periodogram has been suggested by two other authors, Hassler (1993) and Chen *et al.* (1993). In each of these two cases the authors suggest the use of the smoothed periodogram in place of the raw periodogram and investigate the sampling properties of the resulting estimates of the difference parameter via simulation studies. Although these studies overlap to some extent with those reported in the present paper, I should point out that our work was conducted quite independently of that reported in the other papers.

TABLE VI

ARIMA$(1, d, 0)$

	$\phi = 0.1$				$\phi = 0.4$			
	$d = 0.1$	$d = 0.2$	$d = 0.3$	$d = 0.4$	$d = 0.1$	$d = 0.2$	$d = 0.3$	$d = 0.4$
d_{sp}	0.29	0.52	0.74	0.88	0.34	0.58	0.76	0.92
d_p	0.05	0.18	0.32	0.61	0.08	0.18	0.39	0.64

REFERENCES

ANDEL, J. (1986) Long memory time series models. *Kybernetika* 22 (2), 105–23.

ANDERSON, T. W. (1971) *The Statistical Analysis of Time Series*. New York: Wiley.

CHEN, G., ABRAHAM, B. and PEIRIS, S. (1993) Lag window estimation of the degree of differencing in fractionally integrated time series models. Submitted to *J. Time Ser. Anal.*

CHOW, G. C. (1987) *Econometrics*. New York: McGraw-Hill.

FELLER, W. (1951) The asymptotic distribution of the range of sums of independent random variables. *Ann. Math. Stastist.* 22 (4), 427–32.

GEWEKE, J. and PORTER-HUDAK, S. (1983) The estimation and application of long memory time series model. *J. Time Ser. Anal.* 4 (4), 221–38.

GRANGER, C. W. J. and JOYEUX, R. (1980) An introduction to long memory time series models and fractional differencing. *J. Time Ser. Anal.* 1 (1), 15–29.

HASSLER, U. (1993). Regression of spectral estimator with fractionally integrated time series. *J. Time Ser. Anal.* 14, 369–80.

HOSKING, J. (1981) Fractional differencing. *Biometrika* 68 (1), 165–76.

— (1982) *Some Models of Persistence in Time Series Analysis: Theory and Practice 1.* Amsterdam: North-Holland, pp. 641–53.

— (1984) Modelling persistance in hydrological time series using fractional differencing. *Water Resour. Res.* 20 (12) 1898–1908.

HURST, H. E. (1951) Long-term storage capacity of reservoirs. *Trans. Am. Soc. Civil Eng.* 116, 770–99.

— (1956) Methods of using long-term storage in reservoirs. *Proc. Inst. Civil Eng.* 1, 519–43.

KUNSCH, H. R. (1986) Discrimination between monotone trends and long-range dependence. *J. Appl. Prob.* 23, 1025–30.

MANDELBROT, B. B. (1965) Une classe de processus stochastiques homothetiques à soi: Application à la loi climatologique de H. E. Hurst. *C. R. Acad. Sci. Paris* 260, 3274–76.

— and VAN NESS, J. W. (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* 10 (4), 422–37.

— and WALLIS, J. R. (1969a) Computer experiments with fractional Gaussian noises. Part 1, Averages and variances. *Water Resour. Res.* 5 (1), 228–41.

— and — (1969b) Computer experiments with fractional noises. Part 2, Rescaled ranges and spectra. *Water Resour. Res.* 5 (1), 242–59.

— and — (1969c) Computer experiments with fractional Gaussian noises. Part 3, Mathematical appendix. *Water Resour. Res.* 5 (1), 260–67.

MCLEOD, A. I. and HIPEL, K. (1978) Preservation of the rescaled adjusted range. Part 1, A reassessment of the Hurst phenomenon. *Water Resour. Res.* 14 (3), 491–508.

MOOD, A., GRAYBILL, F. and BOES, D. (1974) *Introduction to the Theory of Statistics*. New York: McGraw-Hill.

O'CONELL, P. E. (1974) Stochastic modelling of long-term persistance in streamflow sequences. Ph.D. Thesis, Imperial College London.

PARZEN, E. (1986) *Quantile Spectral Analysis and Long Memory Time Series*. Applied Probability Trust.

PRIESTLEY, M. B. (1981). *Spectral Analysis and Time Series*. London: Academic Press.