

Tests for Hurst effect

BY R. B. DAVIES

Applied Mathematics Division, D.S.I.R., Wellington, New Zealand

AND D. S. HARTE

Biometrics Section, Ministry of Agriculture and Fisheries, Wellington, New Zealand

SUMMARY

We consider the power of tests for distinguishing between fractional Gaussian noise and white noise of a first-order autoregressive process. Our tests are based on the beta-optimal principle (Davies, 1969), local optimality and the rescaled range test.

Some key words: Autoregressive process; Beta-optimal test; Fractional Gaussian noise; Hydrological series; Locally optimal test; Long-term dependence; Rescaled range; Self-similar process; Simulation.

1. INTRODUCTION

Fractional Gaussian noise has been proposed as a model (Mandelbrot & Van Ness, 1968) for the long-term dependence postulated to occur in a variety of hydrological and geophysical time series. See Hurst (1951) for an example based on the flows of the River Nile. A test known as the rescaled range test has been used to detect long-term dependence. Our purpose is to derive optimal tests for detecting long-term dependence and to compare their power with that of the rescaled range test.

Let X_1, \dots, X_n denote a series of observations on which it is proposed to carry out a test for long-term dependence. We suppose that the X_i are normally distributed. Then $\{X_i\}$ is known as fractional Gaussian noise if

$$\text{cov}(X_i, X_j) = \theta \rho(i-j), \quad (1.1)$$

where

$$\rho(i) = \frac{1}{2}|i+1|^{2H} + \frac{1}{2}|i-1|^{2H} - |i|^{2H}. \quad (1.2)$$

The parameter, H , determines the extent of the dependence, $H = \frac{1}{2}$ corresponding to independence and $\frac{1}{2} < H < 1$ corresponding to long-term dependence. Note that, for i large, $\rho(i) \sim H(2H-1)|i|^{2H-2}$ so that the usual asymptotic time series methods (Davies, 1983) may not apply.

First we consider the hypothesis of independence. That is, we wish to test the hypothesis $H = \frac{1}{2}$ against the alternative $\frac{1}{2} < H < 1$. We have derived two tests. The first is the locally optimal test which maximizes the derivative of the power function at $H = \frac{1}{2}$ (Rao, 1973, p. 453). The second is a test introduced by Spjøtvoll (1967) and Davies (1969) which minimizes the value of H for which a prescribed power, say 80%, is reached. Davies (1969) calls this the beta-optimal test and shows how it may be derived as a likelihood ratio test. The formulae for these tests are given in § 3.

The rescaled range test was suggested by Hurst's original observations on the Nile and is based on the minimum size for a reservoir if it were not to overflow or run dry over a given period of time. In § 2 of this paper a version of this test is precisely defined so that its power can be found.

In § 4 we present the results of our power calculations. These show that the locally optimal test is very close to being optimal. However a test which depends only on the first autocorrelation does almost as well, suggesting that long term dependence is not the dominant feature of the fractional Gaussian noise process for the sample sizes we consider. The rescaled range test is found to be substantially less powerful than the other tests requiring about double the sample size to get the same power.

In § 5 we consider tests to discriminate between a first-order autoregressive process and a fractional Gaussian noise. The locally optimal and rescaled range tests turn out to be highly sensitive to the autoregressive alternative. We show that it is very difficult to reliably distinguish these two processes and a series length of at least several hundred will be needed before good discrimination is possible. We look at two tests and estimate their power by simulation.

For the simulations we require a fast way of simulating a process with covariances given by (1.1). Our method is described in the Appendix.

2. THE RESCALED RANGE TEST

The rescaled range test for Hurst effect was introduced as a graphical technique (Mandelbrot & Wallis, 1969). For finding the power of this test we need a precise formulation.

Let $S(t, d)$ be the sample standard deviation of the observations from time t to time $t + d - 1$, and $R(t, d)$ the minimum reservoir size required over the same period if the X_j represent river flows. That is

$$S^2(t, d) = \sum_{j=t}^{t+d-1} X_j^2/d - \{\bar{X}(t, d)\}^2,$$

$$R(t, d) = \max_{1 \leq u \leq d} \left\{ \sum_{j=t}^{t+u-1} X_j - u\bar{X}(t, d) \right\} - \min_{1 \leq u \leq d} \left\{ \sum_{j=t}^{t+u-1} X_j - u\bar{X}(t, d) \right\},$$

where $\bar{X}(t, d) = (X_t + \dots + X_{t+d-1})/d$.

The ratio $R(t, d)/S(t, d)$ is called the rescaled range. The rescaled range test consists of first averaging the values of the rescaled range for a number of values of t for a given value of d . Let $Q(d)$ denote this average. Then $\log \{Q(d)\}$ is plotted against $\log d$. For a fractional Gaussian noise process this graph should have slope H for large d . See Mandelbrot (1975) for more details and Anis & Lloyd (1976) for a calculation of $E\{Q(d)\}$ when the process is white noise.

For our simulations the values of d and t were chosen as follows.

- (i) Choose w values for d : $d_j = \text{int}(n/j)$ for $j = 1, \dots, 6$ or $\text{int}(d_{j-1}/1.15)$ for $j = 7, \dots, w$, where w is chosen so that $d_w = 3$ and int denotes truncation to the integer part of a number.
- (ii) For each value of j choose t_i : $i = 1, \dots, u_j$, where $u_j = \text{int}(n/d_j)$, and $t_i = \text{int}\{\frac{1}{2}(n - d_j u_j)\}$ for $i = 1$, or $t_{i-1} + u_j$ for $i = 2, \dots, u_j$.

Then let

$$Q(d_j) = \sum_{i=1}^{u_j} \{R(t_i, d_j)/S(t_i, d_j)\}/u_j.$$

The slope, \hat{H} , is then found using ordinary least-squares linear regression of $\log \{Q(d_j)\}$ on $\log d_j$. Our test is to reject the hypothesis of white noise in favour of the fractional Gaussian noise alternative if \hat{H} is large. The critical point and power was found by simulation using a sample size of 1000 for calculating the critical point and either 200 or 500 for calculating powers. Results are reported in § 4 of this paper.

3. LOCALLY OPTIMAL AND BETA-OPTIMAL TESTS

We suppose the X_1, \dots, X_n , denoted collectively by the column vector X , are normally distributed with covariance matrix given by (1.1). We wish to test the hypothesis of independence, $H = \frac{1}{2}$, against the alternative $H > \frac{1}{2}$. The locally optimal and beta-optimal tests can be derived in a similar way to the derivations of Spjøtvoll (1967) and Davies (1969).

The locally optimal test has critical region

$$[X'\{c(I - 11'/n) - (I - 11'/n)T(I - 11'/n)\}X < 0], \quad (3.1)$$

where

$$T_{i,j} = \begin{cases} 0 & (i = j), \\ \log 2 & (i = j \pm 1), \\ \frac{1}{2}[|i - j| \log \{1 - (i - j)^{-2}\} + \log \{(|i - j| + 1)/(|i - j| - 1)\}] & \text{otherwise,} \end{cases} \quad (3.2)$$

1 denotes an n -dimensional vector of ones, I denotes the $n \times n$ identity matrix, and c is a constant selected to give the desired significance level. The matrix T is of Toeplitz form and so (3.1) can be evaluated in $O(n \log n)$ operations using a fast Fourier transform program.

The beta-optimal test has critical region

$$[X'\{\Psi_H - c(I - 11'/n)\}X < 0], \quad (3.3)$$

where

$$\Psi_H = \Phi_H^{-1} - \Phi_H^{-1}11'\Phi_H^{-1}/(1'\Phi_H^{-1}1) \quad (3.4)$$

and Φ_H is an $n \times n$ matrix with (i, j) th element equal to $\rho(i - j)$, where ρ is as defined by (1.2). Formula (3.3) involves constants c and H which must be selected as follows. Suppose we wish the power function to reach a power β for as small a value of H in (1.2) as possible. Then c and H must be chosen so that (3.3) has the desired significance level and power equal to β when H in (1.2) is equal to H in (3.3).

4. POWER FUNCTIONS

We report on the powers of the tests introduced in the previous sections. The computer programs made extensive use of subroutines from the IMSL subroutine library for matrix manipulation, eigenvalue calculation and solving nonlinear equations. Distributions of a quadratic form in normal random variables were calculated using a Fortran translation of the algorithm of Davies (1980).

Powers of the beta-optimal, locally optimal, and rescaled range tests for selected values of n and H are displayed in Table 1. In all cases the significance level, $\alpha = 0.05$ and the beta-optimal test uses $\beta = 0.8$. The numbers in parentheses for the rescaled range test are twice the standard error for the estimate of power.

Table 1. Powers of tests against fractional Gaussian noise

n		$H = 0.55$	$H = 0.65$	$H = 0.75$	$H = 0.85$	$H = 0.95$
50	Beta-optimal test	0.13	0.42	0.75	0.93	0.98
	Locally optimal test	0.13	0.42	0.73	0.91	0.98
	Correlation	0.12	0.40	0.73	0.92	0.98
	Rescaled range test	0.13 (0.05)	0.33 (0.07)	0.53 (0.04)	0.77 (0.06)	0.91 (0.04)
100	Beta-optimal test	0.19	0.69	0.96	0.998	1.000
	Locally optimal test	0.19	0.68	0.95	0.996	1.000
	Correlation	0.17	0.65	0.95	0.998	1.000
200	Beta-optimal test	0.29	0.93	0.999	1.000	1.000
	Locally optimal test	0.29	0.92	0.999	1.000	1.000
	Correlation	0.25	0.90	0.999	1.000	1.000
	Rescaled range test	0.17 (0.05)	0.61 (0.07)	0.94 (0.03)	0.98 (0.02)	

For practical purposes the beta-optimal and locally optimal tests have the same power functions and there is no point in using the more complex beta-optimal test. We have also included a test based on the first-order autocorrelation of the process. Perhaps surprisingly, this test does almost as well as the other tests, particularly for low values of n . The rescaled range test is rather less powerful. With $n = 200$, it has a similar power to the other tests with n around 75 to 100.

A second alternative one might wish to look at is to suppose that our process is the sum of a fractional Gaussian noise process with known H , say $H = 0.75$, and white noise. The hypothesis is, as before, that the process is pure white noise. Locally and beta-optimal tests can be found for this situation. In fact, the locally optimal test given in § 3 is close to optimal for $H = 0.75$ and the values of n tested, that is, $n \leq 200$, whereas the locally optimal test derived specifically for this fractional Gaussian noise plus white noise situation was noticeably poorer.

5. AUTOREGRESSIVE HYPOTHESIS

So far we have considered only white noise as the hypothesis. Ideally we would like our tests to be insensitive to the first-order autoregressive alternative

$$X_i = \gamma X_{i-1} + \varepsilon_i, \quad (5.1)$$

with $0 \leq \gamma < 1$ and ε_i a sequence of independent identically distributed normal variables. In Table 2 we have listed the powers of the locally optimal test (3.1) and rescaled range test against a first-order autoregressive alternative. In all cases the powers are unacceptably high although the rescaled range test is more robust than the locally optimal test.

Table 2. Powers of tests against autoregressive alternative

	n	Autoregressive parameter				
		-0.1	0.0	0.1	0.2	0.3
Locally optimal test	50	0.01	0.05	0.15	0.33	0.58
	100	0.01	0.05	0.22	0.53	0.83
	200	0.00	0.05	0.33	0.79	0.98
Rescaled range test	50	0.02 (0.02)	0.05	0.12 (0.04)	0.19 (0.05)	0.32 (0.06)
	200		0.05	0.11 (0.04)	0.27 (0.06)	0.47 (0.07)

We would like to treat (5.1) with $0 \leq \gamma < 1$ as the hypothesis, so that our test has power no more than the significance level, α , under (5.1). For the alternative we could consider pure fractional Gaussian noise or the process generated by (5.1) with the ε_i being fractional Gaussian noise. Most of what follows is applicable to either situation. So far we have not found a fully satisfactory test.

We can find an upper bound on the power of such a test. For a given alternative let B_γ denote the power of the optimum level α test of the hypothesis (5.1) with a given value of γ . Then $\min(B_\gamma; 0 \leq \gamma < 1)$ is an upper bound on the power of the test of the hypothesis (5.1) with unknown γ . Heuristic arguments given by R. B. Davies in his Berkeley thesis suggest that this bound may be quite sharp. For the present problem with a pure fractional Gaussian noise alternative, B_γ can be calculated; the likelihood ratio test statistic being the ratio of two terms of the form $X'\Psi X$ with Ψ being as in (3.4) in the numerator and a corresponding form for an autoregressive process in the denominator.

Our results are given in Table 3. The value of γ at which the minimum is achieved is almost independent of n over the range considered and is listed in the table. As will be seen from the table the powers are very low and a sample of at least several hundred will be required to achieve satisfactory discrimination.

Table 3. Bound on power of test with autoregressive process hypothesis

n	$H = 0.60$	$H = 0.75$	$H = 0.90$
50	0.10	0.18	0.22
100	0.15	0.34	0.45
200	0.25	0.60	0.83
Autoregressive Parameter	0.15	0.36	0.57

One possible test simply looks at the second partial autocorrelation. The test statistic may be derived from a multiple linear regression where one is fitting the first and second lagged X values. The test statistic for the second lagged term is

$$(C_0C_2 - C_1^2) / \{C_0(C_0 - C_2)(C_0^2 + C_0C_2 - 2C_1^2) / n\}^{1/2},$$

where C_0, C_1, C_2 denote the estimators of the variance and first and second autocovariances. Under the hypothesis (5.1) its asymptotic distribution is standard normal and it will tend to be large positive for a fractional Gaussian noise process with $H > \frac{1}{2}$. Note that this test is appropriate only for the pure fractional Gaussian noise alternative as the second partial autocorrelation can be negative in some instances for the combined fractional Gaussian noise autoregressive process.

A second test is derived as a $C(\alpha)$ test. See Neyman (1959) for $C(\alpha)$ tests applied to the independent identically distributed situation. In fact, Neyman's formulae can be applied to more general situations if they are applied to the whole likelihood rather than to the individual observations. Under suitable regularity conditions, for example those of Davies (1983), the resulting tests will be asymptotically optimal. For the present problem the regularity conditions do not apply and asymptotic optimality cannot be guaranteed. Similarly significance probabilities given by $C(\alpha)$ theory may not be correct and must be verified by simulation.

For simplicity suppose we observe $n + 1$ points, X_0, \dots, X_n . Let Z be an n -dimensional vector, composed of the residuals after fitting a first-order autoregressive process: $Z_i = X_i - \gamma X_{i-1}$, where γ is to be replaced by an estimate of the first autocorrelation. The $C(\alpha)$ test statistic is then calculated as a function of the Z_i , regarding X_0 as given.

Denote the average of the Z_i by \bar{Z} . Let $u_i = T_{i,0}$ in (3.2), $v_i(\gamma) = \gamma^{i-1}$,

$$\psi(\gamma) = \frac{\sum_{i=1}^n (n-i)u_i v_i(\gamma)}{\sum_{i=1}^n (n-i)v_i(\gamma)^2}, \quad w_i(\gamma) = u_i - \psi(\gamma)v_i(\gamma),$$

$$D_i = \sum_{j=1}^{n-i} (Z_j - \bar{Z})(Z_{j+i} - \bar{Z}), \quad S = \sum_{i=1}^{n-1} w_i(\gamma)D_i / \left\{ D_0 \sum_{i=1}^{n-1} (n-i)w_i(\gamma)^2 / (n-1) \right\}^{\frac{1}{2}}. \quad (5.2)$$

We must replace γ by an estimate in (5.2). We used the first-order autocorrelation of the X_i . Then our $C(\alpha)$ test rejects the hypothesis when S is large. For $n = 1000$ the normal approximation seems adequate for the distribution of S but fails for $n \leq 200$. In this case an adequate approximation to the critical point was found by setting $\gamma = 0$ in (5.2) and using the quadratic form program.

Simulations were carried out for these two tests and the results are listed in Table 4. In each case 1000 simulations were made. For $H = 0.6, 0.75$, the $C(\alpha)$ test is doing well compared with the partial correlation test and fairly well compared with the bound. When $H = 0.9$ the local nature of the $C(\alpha)$ test is becoming very obvious. The problem may lie in using an estimate for γ that is good only under the hypothesis $H = \frac{1}{2}$. Unfortunately a satisfactory estimate of γ for general H is difficult to find. Nevertheless for $n = 1000$, our $C(\alpha)$ test seems quite reasonable.

Table 4. Power of tests of an autoregressive process hypothesis

Process	$n = 200$		$n = 1000$	
	Part. corr. test	$C(\alpha)$ test	Part. corr. test	$C(\alpha)$ test
White noise	0.03	0.05	0.04	0.04
Autoregressive process				
$\gamma = 0.3$	0.03	0.03	0.04	0.03
0.6	0.04	0.03	0.02	0.03
0.9	0.04	0.11	0.03	0.06
Fractional Gaussian noise				
$H = 0.60$	0.13	0.21	0.39	0.63
0.75	0.31	0.40	0.90	0.98
0.90	0.40	0.12	0.98	0.57

However robustness is claimed only against a first-order autoregressive process and it is hard to see how one could expect to be able to distinguish a fractional Gaussian noise process even from, say, a first-order autoregressive, first-order moving average process without a very long series, say of length 1000.

APPENDIX

Simulating a stationary Gaussian time series

The following method was used to generate the fractional Gaussian noise process used in the simulations. When a fast Fourier program is used for carrying out the transforms $O(n \log n)$ operations are required for generating a series of length $n + 1$.

We wish to generate a Gaussian time series, $\{X_i\}$ for $i = 0, \dots, n$ with autocovariances c_0, c_1, \dots . Let g_k denote the finite Fourier transform of the sequence

$$\{c_0, c_1, \dots, c_{n-1}, c_n, c_{n-1}, \dots, c_1\}.$$

That is

$$g_k = \sum_{j=0}^{n-1} c_j \exp \{2\pi i k j / (2n)\} + \sum_{j=n}^{2n-1} c_{2n-j} \exp \{2\pi i k j / (2n)\} \quad (k = 0, 1, \dots, 2n-1).$$

The series $\{g_k\}$ will be real, but for the method to work it must also be nonnegative. If any of the g_k are negative then the method fails. This has not been a problem in the series used in this paper.

Let $\{Z_k\}$ be a sequence of independent complex normal random variables with independent real and imaginary parts, for $0 \leq k \leq n$, and let $Z_k = \bar{Z}_{2n-k}$ for $n < k < 2n$. Let Z_0 and Z_n be real with variances equal to two, otherwise, for $0 < k < n$, let the real and imaginary parts of Z_k have unit variance.

Then the time series

$$X_j = \frac{1}{2n} \sum_{k=0}^{2n-1} Z_k g_k^{\frac{1}{2}} \exp \{2\pi i j k / (2n)\},$$

for $0 \leq j \leq n$, has the required distribution.

REFERENCES

- ANIS, A. A. & LLOYD, E. H. (1976). The expected value of the adjusted rescaled Hurst range of independent normal summands. *Biometrika* **63**, 111-6.
- DAVIES, R. B. (1969). Beta-optimal tests and an application to the summary evaluation of experiments. *J. R. Statist. Soc. B* **31**, 524-38.
- DAVIES, R. B. (1980). The distribution of a linear combination of chi-squared random variables. Algorithm, AS155. *Appl. Statist.* **29**, 323-33.
- DAVIES, R. B. (1983). Optimal inference in the frequency domain. In *Handbook of Statistics*, **3**, Ed. D. R. Brillinger and P. R. Krishnaiah, pp. 73-92. Amsterdam: North-Holland.
- HURST, H. E. (1951). Long term storage capacity of reservoirs. *Trans. Am. Soc. Civil Eng.* **116**, 770-808.
- MANDELBROT, B. B. (1975). Limit theorems on the self normalized range for weakly and strongly dependent processes. *Z. Wahr. verw. Geb.* **31**, 271-85.
- MANDELBROT, B. B. & VAN NESS, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10**, 422-37.
- MANDELBROT, B. B. & WALLIS, J. R. (1969). Robustness of the rescaled range R/S in the measurement of non-cyclic long run statistical dependence. *Water Resources Res.* **5**, 967-88.
- NEYMAN, J. (1959). Optimal asymptotic tests of composite statistical hypotheses. In *Probability and Statistics, The Harald Cramér Volume*. Ed. U. Grenander, pp. 213-34. New York: Wiley.
- RAO, C. R. (1973). *Linear Statistical Inference and its Applications*, 2nd ed. New York: Wiley.
- SPJØTVOLL, E. (1967). Optimal invariant tests in unbalanced variance components models. *Ann. Math. Statist.* **38**, 422-8.

[Received July 1984. Revised July 1986]