Simulating a Class of Stationary Gaussian Processes Using the Davies-Harte Algorithm, with Application to Long Memory Processes

Peter F. Craigmile

NRCSE

Technical Report Series

NRCSE-TRS No. 057

October 24, 2000

The **NRCSE** was established in 1996 through a cooperative agreement with the United States Environmental Protection Agency which provides the Center's primary funding.

Simulating a class of stationary Gaussian processes using the Davies-Harte algorithm, with application to long memory processes.

Peter F. Craigmile[1].

[1]Department of Statistics, Box 354322, University of Washington, Seattle. WA 98195–4322.

Correspondence Address:
Peter F. Craigmile,
Department of Statistics,
Box 354322,
University of Washington,
Seattle. WA 98195–4322.
pfc@stat.washington.edu

Simulating a class of stationary Gaussian processes using the Davies-Harte algorithm, with application to long memory processes.

Peter F. Craigmile[1]

[1]Department of Statistics, Box 354322, University of Washington, Seattle. WA 98195–4322.

Abstract

We demonstrate that the fast and exact Davies-Harte algorithm is valid for simulating a certain class of stationary Gaussian processes – those with a negative autocovariance sequence for all non-zero lags. The result applies to well known classes of long memory processes – Gaussian fractionally differenced (FD) processes, fractional Gaussian noise (fGn), and the nonstationary fractional Brownian Motion (fBm).

Keywords: Davies-Harte algorithm; Circulant embedding; Simulation; Negative autocovariance sequence; Long memory process; Time series analysis.

1 Introduction

Long memory processes are time series models where the autocovariances decays slowly over increasing time lags. Such phenomena have been observed by experimenters for many years, but it was not until Mandelbrot and van Ness (1968) that statistical models appeared which could accurately represent such processes. It is often important to simulate these processes quickly and exactly (by exactly we mean subject to computer accuracy, and the adequacy of the pseudorandom numbers generated). Being able to simulate in this way allows us to validate models and test hypotheses about the data precisely.

In this paper we investigate the simulation algorithm introduced in the statistical literature by Davies and Harte (1987) and analysed in detail in Wood and Chan (1994). The algorithm is based

on the idea in the engineering community known as circulant embedding (see Johnson (1994) for a good review of this and some other methods of simulation). The Davies-Harte algorithm has a validation step that needs to be verified in each case. When this condition is satisfied, the algorithm is $O(N \log(N))$, where N is the sample size. If this qualification does not hold we might need to resort to a more general $O(N^2)$ algorithm such as one based on the Levinson-Durbin recursions (see e.g., Brockwell and Davis (1991)).

In this paper, we introduce a class of stationary Gaussian processes for which the method always holds – those with negative autocovariance sequences for non-zero lags. In section 2 we review the algorithm, illustrating the validation step. We state our main result in section 3, and the application to the simulation of Gaussian fractionally differenced processes, fractional Gaussian noise, and fractional Brownian motion is given in section 4.

2 The Davies-Harte algorithm

Given we know the autocovariance sequence (ACVS) $\{s_k : k \in \mathbb{Z}\}$ of a stationary Gaussian process, $\{X_t : t \in \mathbb{Z}\}$, the steps to simulate a realisation of length N are:

1. For $k = 0, \ldots, 2N-1$, compute
$$A_{k,N}(s) \equiv \sum_{j=0}^{N} s_j \, e^{-i\pi kj/N} + \sum_{j=N+1}^{2N-1} s_{2N-j} \, e^{-i\pi kj/N};$$

2. Check that $A_{k,N}(s) \geq 0$ for all k;

3. Let $Z_0 \ldots Z_{2N-1}$ be independent mean zero Gaussian random variables with unit variance. Compute the complex valued sequence
$$Y_k \equiv \begin{cases} \sqrt{2NA_{0,N}(s)}\, Z_0, & k = 0; \\ \sqrt{NA_{k,N}(s)}\, (Z_{2k-1} + iZ_{2k}), & 1 \leq k \leq N-1; \\ \sqrt{2NA_{N,N}(s)}\, Z_{2N-1}, & k = N; \\ \sqrt{NA_{k,N}(s)}\, (Z_{4N-1-2k} - iZ_{4N-2k}), & N+1 \leq k \leq 2N-1; \end{cases}$$

4. The simulated realisation of the Gaussian process is obtained as
$$X_t \equiv \frac{1}{2N} \sum_{k=0}^{2N-1} Y_k \, e^{i\pi kt/N}, \qquad t = 0, \ldots, N-1.$$

We can use the fast Fourier transform algorithm to calculate the discrete Fourier transform in steps 1 and 4 and speed up simulation – just round N up to a power of two, and truncate the simulated series at the end.

It is crucial that the nonnegativity condition in step 2 holds. This is not always the case. The autocovariance sequence, $\{\exp(-ck^2) : k \in \mathbb{Z}\}$, with $c = 100$ and $N = 50\ 000$ is one example for which the method does not work (Wood and Chan (1994); Dietrich and Newsam (1997)). It is known that the condition holds if

1. the autocovariance sequence is non-negative, decreasing and convex for non-negative lags (Dietrich and Newsam (1997); Gneiting (1998)); or
2. the autocovariance sequence vanishes for lag $k \geq k_0$ and $N \geq k_0$, for some integer k_0 (Dietrich and Newsam (1997); Chilès and Delfiner (1999); Gneiting (2000a)).

Our key result provides another class of stationary processes, for which the nonnegativity condition is satisfied – namely stationary Gaussian processes with negative autocovariances.

3 The main result

Proposition 3.1 *Suppose $\{X_t : t \in \mathbb{Z}\}$ is a stationary (Gaussian) process with autocovariance sequence (ACVS) $\{s_k : k \in \mathbb{Z}\}$ which satisfies $s_k < 0$ for $k \neq 0$. Then $A_{k,N}(s) \geq 0$ for all $k = 0, \ldots, 2N-1$ and $N \geq 1$.*

Proof By rearranging the terms in the sum and using some algebra we have

$$A_{k,N}(s) \quad = \quad s_0 + (-1)^k s_N + 2\sum_{j=1}^{N-1} s_j \cos(\pi kj/N).$$

Since $s_k < 0$ and both $(-1)^k$ and $\cos(\pi kj/N)$ are bounded above by 1

$$s_0 + (-1)^k s_N + 2\sum_{j=1}^{N-1} s_j \cos(\pi kj/N) \quad \geq \quad s_0 + s_N + 2\sum_{j=1}^{N-1} s_j \quad = \quad A_{0,N}(s).$$

Thus $A_{0,N}(s)$ is a lower bound for $A_{k,N}(s)$, and it suffices to prove that $A_{0,N}(s) \geq 0$ for all $N \geq 1$.

We consider two cases.

Case (i) – Suppose that $\{s_k\}$ is absolutely summable. By Corollary 4.3.2 on p. 120 of Brockwell and Davis (1991) the spectral density function, $S(f)$, of the process, given by

$$S(f) \;=\; \sum_{k\in\mathbb{Z}} s_k\, e^{-i2\pi fk},$$

satisfies

$$S_X(0) \;=\; \sum_{k\in\mathbb{Z}} s_k \;\geq\; 0.$$

Now $A_{0,N}(s)$ decreases monotonically with N to $S_X(0) \geq 0$, and the result follows.

Case (ii) – In the general case, consider

$$s_k^{(\lambda)} \;\equiv\; \lambda^{|k|} s_k.$$

for $\lambda \in (0,1)$. Note that $\lambda^{|k|}$ is the ACVS of an AR(1) process. We then have that $s_k^{(\lambda)}$ is a valid ACVS since it is the ACVS of the product of $\{X_t\}$ and an indepedent AR(1) process. Because λ is positive, $s_k^{(\lambda)} < 0$ for $k \neq 0$. The sequence $\{\lambda^{|k|}\}$ is absolutely summable and hence so is $s_k^{(\lambda)}$, thus by case (i)

$$A_{0,N}(s^{(\lambda)}) \geq 0 \tag{1}$$

for all $\lambda \in (0,1)$ and $N \geq 1$. Letting $\lambda \to 1$ in equation (1) we find that $A_{0,N}(s)$ is non-negative, as desired.

4 Three classes of long memory processes

In this section we apply our result, along with some others in the literature, to the simulation of Gaussian fractionally differenced (FD) processes, fractional Gaussian noise (fGn), and fractional Brownian motion (fBm). Each include examples of long memory processes, i.e., there exists a real number $\alpha \in (0,1)$ and constant K such that the spectrum of the process, S(f), satisfies

$$\lim_{f\to 0} \frac{S(f)}{K|f|^{-\alpha}} = 1.$$

Basically this amounts to saying that the spectral density function has a pole at zero.

4.1 Gaussian fractionally differenced processes

The FD process is a long memory dependence model which has become popular in recent years, mainly due to its tractable mathematical properties. The process was originally proposed by Granger and Joyeux (1980) and Hosking (1981) as an extension to ARIMA$(0, d, 0)$ models to allow for fractional values of d. The partial autocorrelations for this process are given in Hosking (1981), and thus $O(N^2)$ simulation via the Levinson-Durbin recursions is possible – we can do better with the Davies-Harte algorithm.

For $d \in [-1/2, 1/2)$ and $\sigma_\epsilon^2 > 0$, the Gaussian process $\{X_t : t \in \mathbb{Z}\}$ is a $FD(d, \sigma_\epsilon^2)$ (or ARFIMA$(0, \mathrm{d}, 0)$) process if it has a spectral density function

$$S(f) = \sigma_\epsilon^2 |2\sin(\pi f)|^{-2d}, \qquad |f| \leq 1/2. \tag{2}$$

Here d is known as the *fractional difference parameter* and σ_ϵ^2 is the *innovation variance*. For $d \in (-1/2, 1/2)$ the process is stationary and invertible, and is a white noise (i.e., uncorrelated) process for $d = 0$. For $d = -1/2$ the process is stationary, but noninvertible. We can extend this class of models by letting $d \geq 1/2$ in equation (2), and obtain a class of non-stationary processes which become stationary after differencing $\lfloor d + 1/2 \rfloor$ times. Differencing we can let $d < -1/2$ to obtain a stationary, but noninvertible, process. For $d \in [-1/2, 1/2)$ the covariance sequence can be shown to be (Beran (1994), Hosking (1981) for the $d = -1/2$ case)

$$s_k = \sigma_\epsilon^2 \frac{(-1)^k \, \Gamma(1 - 2d)}{\Gamma(1 - d + k) \, \Gamma(1 - d - k)}. \tag{3}$$

For $k > 0$ define $a_k(d) \equiv (k - 1 + d)/(k - d)$. A more computationally efficient definition is then given by

$$s_k = \begin{cases} \sigma_\epsilon^2 \dfrac{\Gamma(1 - 2d)}{\Gamma^2(1 - d)}, & k = 0; \\ s_{k-1} \, a_k(d), & k \geq 1. \end{cases} \tag{4}$$

Lemma 4.1

$$a_k(d) = \begin{cases} > 0, & d \in (0, 1/2), \ k \geq 1; \\ > 0, & d \in [-1/2, 0), \ k = 1; \\ < 0, & d \in [-1/2, 0), \ k > 1. \end{cases}$$

Thus

$$s_k = \begin{cases} > 0, & d \in (0, 1/2),\ k \geq 0; \\ > 0, & d \in [-1/2, 0),\ k = 0; \\ < 0, & d \in [-1/2, 0),\ k \geq 1. \end{cases}$$

Proof Follows directly from the definition of $a_k(d)$ and equation (4).

Lemma 4.2 *The Davies-Harte algorithm applies to simulating a Gaussian FD(d,σ_ϵ^2) process for* $d \in [-1/2, 1/2)$.

Proof For $d \in [-1/2, 0)$, the result is immediate from Lemma 4.1 and Proposition 3.1. For $d \in [0, 1/2)$ we need to appeal to Dietrich and Newsam (1997), Theorem 2 or Gneiting (1998), Theorem 1. As stated earlier a sufficient condition is that the autocovariance sequence is non-negative, decreasing and convex. Non-negativity follows from Lemma 4.1 and the decreasing nature follows from equation (4) and the fact that $a_k(d)$ is decreasing with increasing k. To show convexity it is sufficient to prove that

$$s_k - 2s_{k+1} + s_{k+2} \geq 0,$$

for each integer $k \geq 0$. Thus

$$\begin{aligned} s_k - 2s_{k+1} + s_{k+2} &= s_k \left[1 - 2a_k(d) + a_{k+1}(d)a_k(d)\right] \\ &= \frac{2(2k^2 + d - 1)}{(k - d)(k + 1 - d)} \geq 0, \end{aligned}$$

when $d \in [0, 1/2)$, and the result follows.

The above lemma proves a conjecture of Caccia et al. (1997), p. 612. Example realisations of FD processes are show for various values of d in Figure 1. We can simulate Gaussian FD(d, σ_ϵ^2) for $d \geq 1/2$ by cumulatively summing a FD($d - \lfloor d + 1/2 \rfloor, \sigma_\epsilon^2$) process $\lfloor d + 1/2 \rfloor$ times. To simulate an ARFIMA(p, d, q) model with non-zero regressive and/or moving average components we use standard ARIMA simulation methods based on the FD process innovations (see Brockwell and Davis (1991) for example).

Further properties along with an extensive history of the FD process can be found in Beran (1994) and Samorodnitsky and Taqqu (1994), sections 7.13 and 14.7.

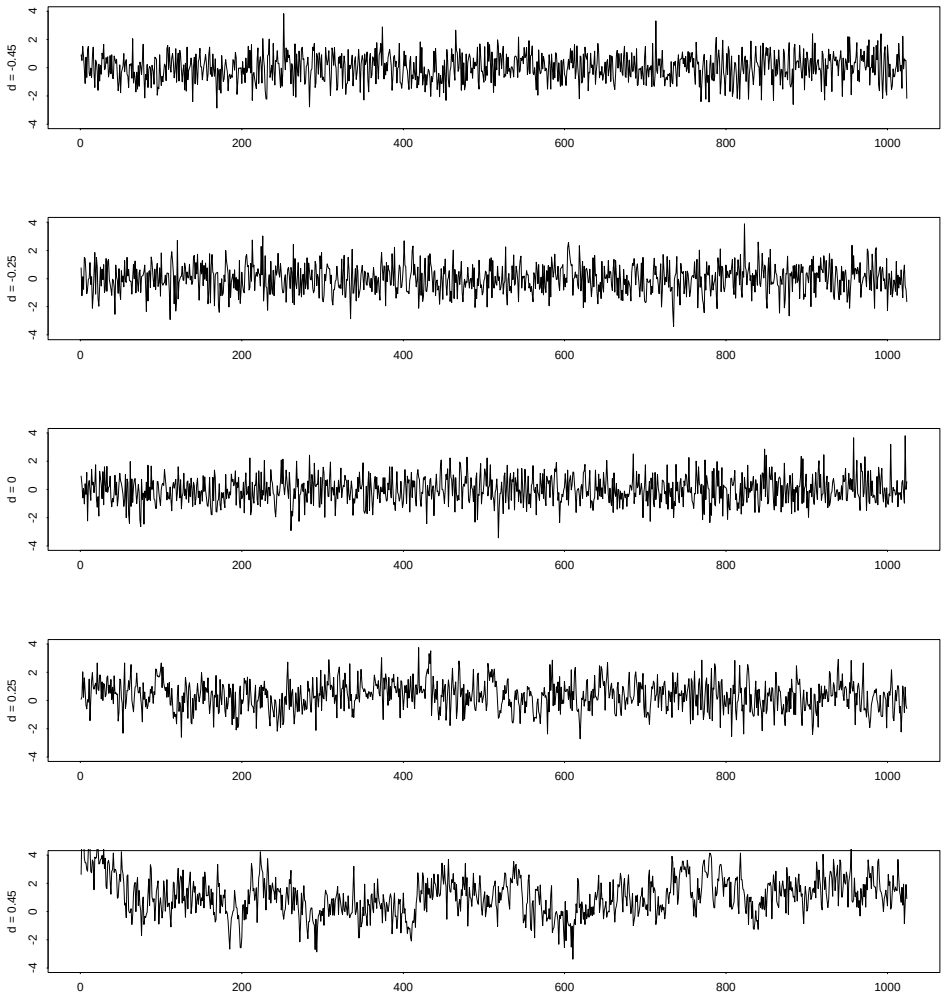

Figure 1: Example realisations of a FD(d) process for various values of d.

4.2 Fractional Gaussian noise

For $0 < H < 1$, fGn(H) is a mean zero stationary Gaussian process $\{X_k : k \in \mathbb{Z}\}$ with autocovariance sequence

$$s_k = \frac{\sigma_0^2}{2}\left[|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}\right]$$

(one can also define this process in terms of fractional Brownian motion – see the next section). Here H is known as the *Hurst parameter*, and σ_0^2 is the *innovation variance*. Letting $D_H \equiv H\,\Gamma(2H)\,\sin(\pi H)\,2^{-2H}\,\pi^{-2H-2}$, the spectral density function is given by

$$S(f) = D_H\,\sigma_0^2\,\sin^2(\pi f)\sum_{k\in\mathbb{Z}}|f+k|^{-(2H+1)}, \qquad f \in [-1/2, 1/2].$$

The long memory behaviour of the process is demonstrated by seeing that $s_k \sim \sigma_0^2 H(2H-1)k^{2H-2}$ $(k \to \infty)$ and $S(f) \sim 2\pi^2\sigma_0^2 D_H |f|^{1-2H}$ $(H \neq 1/2,\ f \to 0)$.

We now apply our result to show that we can use the Davies-Harte algorithm to simulate fractional Gaussian noise with $0 < H < 1/2$ (Chilès and Delfiner (1999) and Gneiting (2000b) show that the algorithm is valid for $1/2 \leq H < 1$).

Lemma 4.3 *For $0 < H < 1/2$ and $k \neq 0$, $s_k < 0$.*

Proof Let $\alpha \equiv 2H$. Then $0 < H < 1/2$ is equivalent to $0 < \alpha < 1$. When $k = 1$

$$\frac{2s_1}{\sigma_0^2} = 2^\alpha - 2,$$

which is negative for $0 < \alpha < 1$, and for $k > 1$

$$\begin{aligned}
\frac{2s_k}{\sigma_0^2} &= (k+1)^\alpha - 2k^\alpha + (k-1)^\alpha \\
&= k^\alpha\left[(1+1/k)^\alpha + (1-1/k)^\alpha - 2\right] \\
&= k^\alpha\sum_{r=1}^{\infty}\binom{\alpha}{r}\left[k^{-r} + (-k)^{-r}\right] \\
&= 2k^\alpha\sum_{r=1}^{\infty}\binom{\alpha}{2r}k^{-2r} \\
&= 2k^\alpha\sum_{r=1}^{\infty}\frac{\Gamma(\alpha+1)\,k^{-2r}}{\Gamma(2r+1)\Gamma(\alpha-2r+1)},
\end{aligned}$$

which again is negative since $\Gamma(\alpha - 2r + 1)$ is negative. By symmetry of the covariance function the result holds for negative k.

Thus by Proposition 3.1 the Davies-Harte algorithm applies to the simulation of fGn processes, with $0 < H < 1/2$. Some example realisations of fGn processes for various values of H are shown in Figure 2. Further properties along with an extensive history of this process can be found in Samorodnitsky and Taqqu (1994), sections 7.2 and 14.7.

4.3 Fractional Brownian motion

The fBm, $\{\mathbb{B}_H(t) : t \in \mathbb{R},\ 0 < H < 1\}$, of Mandelbrot and van Ness (1968) is a generalisation of the standard Brownian motion defined by $\mathbb{B}_H(0) = 0$ and

$$\begin{aligned}\mathbb{B}_H(t) - \mathbb{B}_H(0) &= \frac{1}{\Gamma(H+1/2)} \left\{ \int_{-\infty}^{0} \left[(t-s)^{H-1/2} - (-s)^{H-1/2} \right] d\mathbb{B}(s) \right. \\ &\qquad \left. + \int_0^t (t-s)^{H-1/2}\, d\mathbb{B}(s) \right\},\end{aligned}$$

where $\mathbb{B}(s)$ is a standard Brownian motion. This process is self-similar: for $a > 0$

$$\mathbb{B}_H(at) \quad =_d \quad a^H \mathbb{B}_H(t),$$

and taking differences

$$X_k \quad \equiv \quad \mathbb{B}_H(k+1) - \mathbb{B}_H(k) \qquad k \in \mathbb{Z},$$

where $\{X_k : k \in \mathbb{Z}\}$ is a fGn(H) process. To produce a realisation of a $\mathbb{B}_H(H)$ process on a finite grid of points we simulate a realisation of an fGn(H) process using the Davies-Harte algorithm, cumulatively sum these innovations, and use the self-similarity relation to scale according to the sampling rate (see Chilès and Delfiner (1999) for more details). More information on fBm can be found in Samorodnitsky and Taqqu (1994), section 7.2.

Acknowledgements

The author gratefully acknowledges support for this research from an STTR grant from AFOSR (MathSoft, Inc., and the University of Washington), and an EPA grant (National Research Center

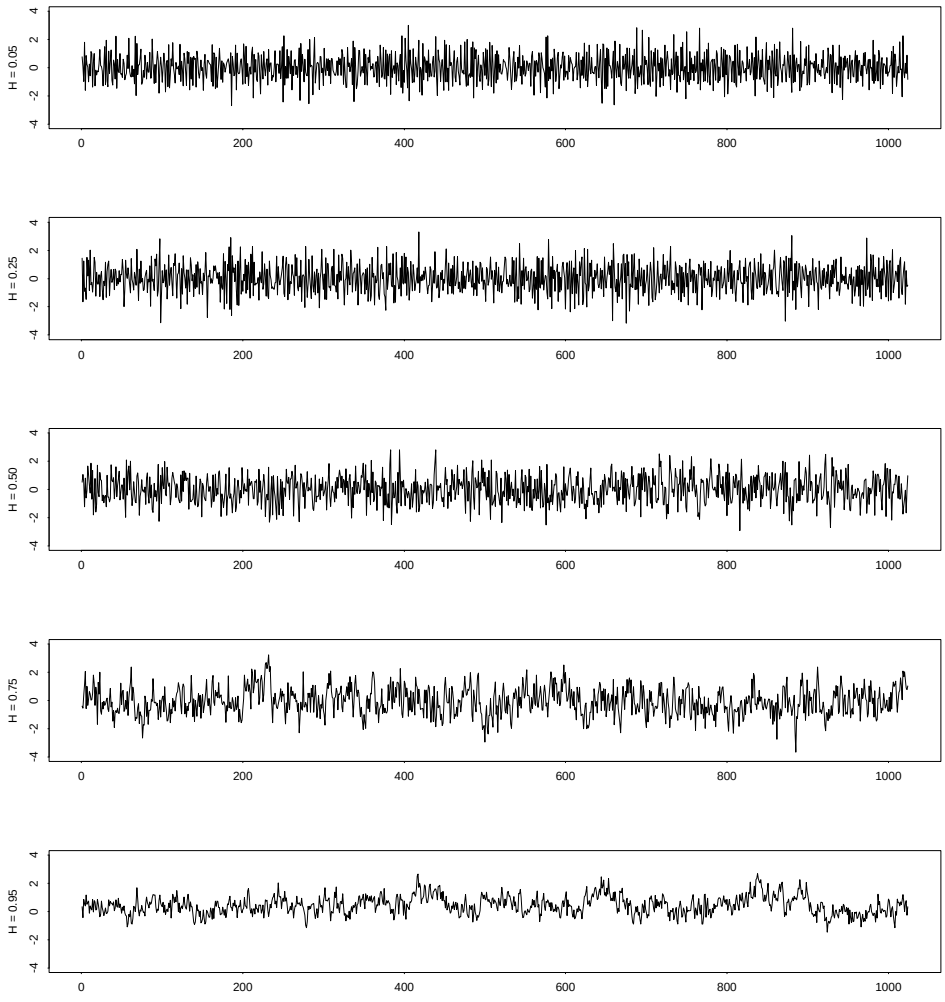

Figure 2: Example realisations of a FGN(H) process for various values of H.

for Statistics and the Environment). He would also like to thank Dr. Tilmann Gneiting and Dr. Donald B. Percival for their many useful comments and advice.

References

Beran, J. (1994). *Statistics for Long Memory Processes*, Volume 61 of *Monographs on Statistics and Applied Probability*. New York: Chapman and Hall.

Brockwell, P. J. and R. A. Davis (1991). *Time Series. Theory and Methods (Second Edition)*. New York: Springer-Verlag.

Caccia, D. C., D. Percival, M. J. Cannon, G. Raymond, and J. B. Bassingthwaighte (1997). Analyzing exact fractal time series: evaluating dispersional analysis and rescaled range methods. *Physica A 246*, 609–632.

Chilès, J. and P. Delfiner (1999). *Geostatistics: modeling spatial uncertainty*. New York: Wiley.

Davies, R. B. and D. S. Harte (1987). Tests for Hurst effect. *Biometrika 74*, 95–101.

Dietrich, C. R. and G. N. Newsam (1997). Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. *SIAM Journal on Scientific and Statistical Computing 18*, 1088–1107.

Gneiting, T. (1998). Simple tests for the validity of correlation function models on the circle. *Statistics and Probability Letters 39*, 119–122.

Gneiting, T. (2000a). Compactly supported correlation functions. Technical Report 45, National Research Center for Statistics and the Environment, University of Washington.

Gneiting, T. (2000b). Power-law correlations, related models for long-range dependence, and their fast simulation. *To appear in Journal of Applied Probability*.

Granger, C. W. J. and R. Joyeux (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis 1*, 15–29.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika 68*(1), 165–176.

Johnson, G. E. (1994). Constructions of particular random processes. *Proceedings of the IEEE 82*(2), 270–284.

Mandelbrot, B. and J. van Ness (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review 10*(4), 422–436.

Samorodnitsky, G. and M. S. Taqqu (1994). *Stable Non-Gaussian Random Processes: Stochastic Models With Infinite Variance*. New York: Chapman & Hall.

Wood, A. T. A. and G. Chan (1994). Simulation of stationary Gaussian processes in $[0, 1]^d$. *Journal of Computational and Graphical Statistics 3*, 409–432.