Exact Simulation of Time-Varying Fractionally Differenced Processes

Donald B. Percival[1,2] and William L. B. Constantine[2]

[1]Applied Physics Laboratory, Box 355640, University of Washington, Seattle, WA 98195–5640, USA

[2]Insightful, Inc., 1700 Westlake Avenue North, Suite 500, Seattle, WA 98109–3044, USA

Correspondence to Dr. Donald B. Percival, Insightful, Inc., 1700 Westlake Avenue North, Suite 500, Seattle, Washington 98109–3044, USA; Phone: (206)–802–2258; Fax: (206)–283–8691; E-mail: `dbp@insightful.com`

Abstract

Time-varying fractionally differenced (TVFD) processes can serve as useful models for certain time series whose statistical properties evolve over time. The spectral density function for a TVFD process obeys a power law whose exponent can be time dependent. In contrast to locally stationary or locally self-similar processes, the power law exponent for a TVFD process is not restricted to certain intervals, which is of practical importance for modeling time series of, e.g., atmospheric turbulence. In this paper we construct a uniform representation for Gaussian TVFD processes that allows the power law exponent to evolve in an arbitrary manner. Even though this representation in general involves a time-dependent linear combination of an infinite number of random variables from a Gaussian white noise process, we demonstrate that simulations with exactly correct statistical properties can be achieved based upon two well-known approaches, each of which involves a finite portion of a white noise process. The first approach is based on the modified Cholesky decomposition, and the second, on circulant embedding. For either approach, the resulting algorithm for generating simulations of a TVFD process can be simply described as 'cutting and pasting' pieces of simulations from several different FD processes, all created from a single realization of a white noise process. Use of these exact methods will ensure that Monte Carlo studies of the statistical properties of estimators for TVFD processes are not adversely influenced by imperfections arising from the use of approximate simulation methods.

Keywords: Circulant Embedding; Gaussian Process; Locally Self-Similar Process; Locally Stationary Process; Power Law Process; Time Series Analysis

1 Introduction

Let $\{X_t : t \in \mathbb{Z}^*\}$ be a discrete parameter Gaussian (i.e., normally distributed) stochastic process with zero mean, where $\mathbb{Z}^* = \{0, 1, 2, \ldots\}$ is the set of all nonnegative integers. We say that $\{X_t\}$ obeys a power law at low frequencies if it possesses a spectral density function (SDF) $S_X(\cdot)$ such that

$$\lim_{f \to 0} \frac{S_X(f)}{C|f|^\alpha} = 1, \tag{1}$$

where α can be any real-valued number, while $C > 0$; i.e., $S_X(f) \approx C|f|^\alpha$ at low frequencies. Well-known examples of such a process are fractional Gaussian noise, for which $-1 < \alpha < 1$, and discrete parameter fractional Brownian motion, for which $-3 < \alpha < -1$ (Mandelbrot and van Ness, 1968; Sinai, 1976; Flandrin, 1989; Beran, 1994). When $\alpha > -1$, a process satisfying Equation (1) is stationary; when $\alpha \leq -1$, the process is nonstationary, but its backward difference of order $d \equiv \lfloor (1-\alpha)/2 \rfloor$, i.e.,

$$X_t^{(d)} \equiv (1-B)^d X_t \equiv \sum_{k=0}^{d} \binom{d}{k} (-1)^k X_{t-k}, \quad t \in \mathbb{Z}^*, \tag{2}$$

is a stationary process. In the above, $\lfloor x \rfloor$ is the largest integer less than or equal to x; B is the backward shift operator (hence $(1-B)X_t = X_t - X_{t-1}$, $B^2 X_t = X_{t-2}$ and so forth); and we take X_t to be zero when $t \leq -1$. When $-1 < \alpha < 0$, Equation (1) is a viable definition for a stationary long memory process (Beran, 1994, p. 42).

Processes satisfying Equation (1) have proven to be useful as models for time series collected in areas such as geophysics, oceanography, atmospheric science, astronomy, economics and electrical engineering; however, they suffer from the limitation that their statistical properties are invariant over time and hence cannot suitably describe certain physical phenomena. Because of this limitation, there has been considerable recent interest in processes that locally obey a power law; i.e., we assume that $\{X_t\}$ possesses a time-varying

SDF such that $S_{X_t}(f) \approx C_t|f|^{\alpha_t}$ over a certain range of frequencies, where the level C_t and the power law exponent α_t are now allowed to change over time. Two processes that fit into this framework are locally stationary processes (Priestley, 1965; Dahlhaus, 1996; Dahlhaus, 1997; Mallat *et al.*, 1998; Carmona *et al.*, 1998; Mallat, 1999) and locally self-similar processes (Gonçalvès and Flandrin, 1993; Flandrin and Gonçalvès, 1994; Wang *et al.*, 2001). These two classes of processes have served as successful models for some time series, but both have the disadvantage of restricting α_t to be in a certain range; i.e., locally stationary and locally self-similar processes require, respectively, $\alpha_t > -1$ and $-3 < \alpha_t < -1$ for all t. These restrictions are not reasonable for certain time series. For example, in fitting locally self-similar processes to a geophysical time series, Wang *et al.*, 2001, obtained estimates of α_t outside of the allowable range $(-3, -1)$ and commented that these estimates were not compatible with the assumed model. In a recent analysis of an aerothermal turbulence time series, Constantine *et al.*, 2001, found evidence that α_t varied between 0 and -2, a range that cannot be fully achieved by either locally stationary or locally self-similar processes alone.

To circumvent restricting the range of α_t, we consider in this paper a time-varying fractionally differenced (TVFD) process. Fractionally differenced (FD) processes were introduced by Granger and Joyeux, 1980, and Hosking, 1981, and obey Equation (1) with α allowed to be any real-valued number. Thus, in contrast to locally stationary or locally self-similar processes, a TVFD process has no restrictions on α_t. The definition of a TVFD process generally involves a moving average of an infinite number of random variables (RVs) from a Gaussian white noise process $\{\epsilon_t\}$ with zero mean and unit variance. Because it involves an infinite sum, this definition cannot be directly used for generating exact realizations from TVFD processes. We consider two equivalent formulations for a TVFD, both of which involve only a finite number of RVs from $\{\epsilon_t\}$. Both formulations permit the generation of exact realizations, thus ensuring that, if we conduct a Monte Carlo experiment to determine the properties of a statistic formed from a TVFD process, our results will reflect the correct properties of the statistic rather than inaccuracies in the simulation process (see Wood and Chan, 1994, for further discussion on the meaning and importance of exact realizations).

The remainder of this paper is organized as follows. In §2 we review the definition of an FD process, after which we define a TVFD process in §3 and consider two methods for generating exact realizations from such a process (§3.1 and §3.2). We note in §3.3 that both methods can be described very simply as 'cutting and pasting' together simulations from different FD processes, each created from the same set of Gaussian white noise deviates. We then give some examples of realizations from TVFD processes in §4, followed by some concluding comments in §5.

2 Definition of Fractionally Differenced Processes

As before, let $\{X_t : t \in \mathbb{Z}^*\}$ be a discrete parameter Gaussian stochastic process with zero mean. We say that $\{X_t\}$ is a stationary FD process if it has an SDF given by

$$S_X(f) \equiv \frac{\sigma_\epsilon^2}{[4\sin^2(\pi f)]^\delta}, \quad |f| \leq \frac{1}{2}, \tag{3}$$

where $\sigma_\epsilon^2 > 0$ and $\delta < 1/2$ are the FD parameters. The parameter δ determines the shape of the SDF, while σ_ϵ^2 merely controls its level. Using the small angle approximation $\sin(x) \approx x$, we see that

$$S_X(f) \approx \frac{\sigma_\epsilon^2}{|2\pi f|^{2\delta}}$$

at low frequencies, so an FD process obeys Equation (1) with $\alpha = -2\delta$. When $\delta = 0$, we have $S_X(f) = \sigma_\epsilon^2$, which is the SDF for a white noise process. When $\delta \neq 0, -1, -2, \ldots$, the autocovariance sequence (ACVS) $\{s_{X,\tau}\}$ for an FD process is given by

$$s_{X,\tau} = \sigma_\epsilon^2 \frac{\sin(\pi\delta)\Gamma(1-2\delta)\Gamma(\tau+\delta)}{\pi\Gamma(1+\tau-\delta)}, \quad \tau \in \mathbb{Z}, \tag{4}$$

where $\mathbb{Z}$ is the set of all integers; when $\delta = 0, -1, -2, \ldots$, the ACVS is given by

$$s_{X,\tau} = \sigma_\epsilon^2 \frac{(-1)^\tau \Gamma(1-2\delta)}{\Gamma(1+\tau-\delta)\Gamma(1-\tau-\delta)}, \quad 0 \le |\tau| \le -\delta, \tag{5}$$

whereas $s_{X,\tau} = 0$ for $|\tau| > -\delta$. For all $\delta < 1/2$, we can write

$$s_{X,0} = \sigma_\epsilon^2 \frac{\Gamma(1-2\delta)}{\Gamma^2(1-\delta)}, \tag{6}$$

while the remainder terms of its ACVS can be computed using the recursion

$$s_{X,\tau} = s_{X,\tau-1} \frac{\tau+\delta-1}{\tau-\delta}, \qquad \tau = 1, 2, \ldots \tag{7}$$

(for negative lags τ, recall that $s_{X,-\tau} = s_{X,\tau}$). For any $\delta < 1/2$, the corresponding partial autocorrelation sequence (PACS) is given by

$$\phi_{t,t} \equiv \frac{\delta}{t-\delta}, \quad t \in \mathbb{Z}^+, \tag{8}$$

where $\mathbb{Z}^+ \equiv \{1, 2, \ldots\}$ (the results above are discussed in Granger and Joyeux, 1980, for the case $-1 < \delta < 1/2$ and in Hosking, 1981, for the case $-1/2 < \delta < 1/2$; see the appendix for a discussion on the extension of these previous results to all $\delta < 1/2$).

For $\delta \ge 1/2$ we say that $\{X_t : t \in \mathbb{Z}^*\}$ is a nonstationary FD process if the dth order backward difference $\{X_t^{(d)}\}$ of Equation (2) is a stationary FD process with parameter $\delta^{(s)} \equiv \delta - d$, where $d \equiv \lfloor \delta + 1/2 \rfloor \in \mathbb{Z}^+$, and $-1/2 \le \delta^{(s)} < 1/2$. Since the squared gain function for a first order backward difference is given by $4\sin^2(\pi f)$, we can use the theory of linear filters to argue that the SDF for a nonstationary FD process is given by Equation (3) (Yaglom, 1958). In what follows, we make the further assumption that $\{X_t\}$ can be expressed in terms of a stationary process $\{X_t^{(d)}\}$ via d successive cumulative summations; i.e., with $X_t^{(0)} \equiv X_t$ we can write

$$X_t^{(k-1)} \equiv \sum_{l=0}^{t} X_l^{(k)}, \quad k = d, d-1, \ldots, 1. \tag{9}$$

For example, when $d = 2$, we start with the stationary FD process $\{X_t^{(2)}\}$ with parameter $\delta^{(s)}$ and cumulatively sum it to obtain $\{X_t^{(1)}\}$, which in turn we sum to get the nonstationary FD process $\{X_t\}$ with parameter $\delta = \delta^{(s)} + 2$:

$$X_t^{(1)} = \sum_{l=0}^{t} X_l^{(2)} \text{ and } X_t = \sum_{l=0}^{t} X_l^{(1)}, \quad t \in \mathbb{Z}^*. \tag{10}$$

This manner of expressing $\{X_t\}$ in terms of its associated stationary process $\{X_t^{(d)}\}$ is consistent with Equation (2) if we recall that X_t is defined to be zero for all $t \leq -1$.

3 Definition and Simulation of TVFD Processes

If $\{X_t : t \in \mathbb{Z}^*\}$ is a stationary zero mean Gaussian FD process (i.e., $\delta < 1/2$), we can express the process as an infinite weighted moving average of a Gaussian white noise process $\{\epsilon_t : t \in \mathbb{Z}\}$ with zero mean and unit variance, namely,

$$X_t = \sum_{k=0}^{\infty} \sigma_\epsilon \frac{(-1)^k \Gamma(1-\delta)}{\Gamma(k+1)\Gamma(1-\delta-k)} \epsilon_{t-k} \equiv \sigma_\epsilon \sum_{k=0}^{\infty} a_k(\delta) \epsilon_{t-k} \tag{11}$$

(Granger and Joyeux, 1980, and Hosking, 1981). By cumulatively summing a stationary FD process an appropriate number of times, we can obtain a corresponding representation for a nonstationary FD process; however, the weights in the moving average will now necessarily depend on t. For example, suppose $1/2 \leq \delta < 3/2$, and suppose we use $\delta^{(s)} = \delta - 1$ in place of δ on the right-hand side of Equation (11) to form the stationary FD process $\{X_t^{(1)}\}$. If we then cumulatively sum this process to form a nonstationary FD process $\{X_t\}$ with parameter δ, we obtain

$$X_t = \sum_{l=0}^{t} X_l^{(1)} = \sigma_\epsilon \sum_{l=0}^{t} \sum_{k=0}^{\infty} a_k(\delta - 1)\epsilon_{l-k}.$$

If we let

$$a_{t,k}(\delta) \equiv \sum_{l=\max\{0,k-t\}}^{k} a_k(\delta - 1), \quad \frac{1}{2} \leq \delta < \frac{3}{2},$$

we can now write

$$X_t = \sigma_\epsilon \sum_{k=0}^{\infty} a_{t,k}(\delta)\epsilon_{t-k}. \tag{12}$$

For FD processes with $\delta \geq 3/2$, we can obtain a representation similar to the above if we define $a_{t,k}(\delta)$ appropriately. If we further define $a_{t,k}(\delta) = a_k(\delta)$ when $\delta < 1/2$, then Equation (12) gives us a representation for all FD processes.

We can define a time-varying FD process by allowing the parameters δ and σ_ϵ^2 in Equation (12) to change over time; i.e., we define such a process via

$$X_t(\boldsymbol{\beta}_t) \equiv \sigma_{\epsilon,t} \sum_{k=0}^{\infty} a_{t,k}(\delta_t)\epsilon_{t-k}. \tag{13}$$

where $\boldsymbol{\beta}_t \equiv [\delta_t, \sigma_{\epsilon,t}^2]^T$. If we make the restriction $\delta_t < 1/2$ for all t, then $\{X_t(\boldsymbol{\beta}_t)\}$ can be an example of a locally stationary process (Whitcher and Jensen, 2000); if we require $1/2 < \delta_t < 3/2$ for all t, it can be an example of a locally self-similar process (Wang *et al.*, 2001).

Let us now turn to the question of how to generate a realization from a finite portion $X_0(\boldsymbol{\beta}_0), X_1(\boldsymbol{\beta}_1), \ldots, X_{N-1}(\boldsymbol{\beta}_{N-1})$ of a TVFD process. Direct use of Equation (13) poses a problem because of the infinite summation. One strategy would be to truncate the summation to, say, M terms, yielding

$$X_t(\boldsymbol{\beta}_t) \approx \sigma_{\epsilon,t} \sum_{k=0}^{M-1} a_{t,k}(\delta_t)\epsilon_{t-k};$$

however, it is unclear what choice of M yields a sufficiently accurate approximation (this is particularly true when δ_t is slightly less than 1/2 because of the slow decay of the weights). Rather than seeking an approximation based upon Equation (13), we instead consider two well-known alternative schemes that allow us to generate exact realizations. Both schemes are based upon forming linear combinations of a finite number of RVs from a white noise process such that the resulting RVs have the same multivariate distribution as $X_0(\boldsymbol{\beta}_0), X_1(\boldsymbol{\beta}_1), \ldots, X_{N-1}(\boldsymbol{\beta}_{N-1})$. The first scheme is based on a modified Cholesky decomposition (§3.1) and yields RVs of the form

$$\sigma_{\epsilon,t} \sum_{k=0}^{t} b_{t,k}(\delta_t)\epsilon_k, \qquad t = 0, 1, \ldots, N-1;$$

the second is a circulant embedding scheme (§3.2) and yields

$$\sigma_{\epsilon,t} \sum_{k=0}^{2N-1} c_{t,k}(\delta_t)\epsilon_k, \qquad t = 0, 1, \ldots, N-1.$$

3.1 Modified Cholesky Decomposition Approach

Let us first review the use of the modified Cholesky decomposition for generating a time series of length N that can be regarded as a realization from a portion $\mathbf{X} \equiv [X_0, X_1, \ldots, X_{N-1}]^T$ of a stationary FD process with parameters $\delta < 1/2$ and $\sigma_\epsilon^2 > 0$ (Hosking, 1984). Given the PACS of Equation (8), we can recursively compute

$$\phi_{t,k} \equiv \phi_{t-1,k} - \phi_{t,t}\phi_{t-1,t-k} \quad \text{for } t = 2, \ldots, N-1 \text{ and } k = 1, \ldots, t-1.$$

The coefficients $\{\phi_{t,k} : k = 1, \ldots, t\}$ can be used to form the best linear predictor $\widehat{X}_t$ of X_t, given $X_{t-1}, \ldots, X_0$; i.e., the predictor

$$\widehat{X}_t \equiv \sum_{k=1}^{t} \phi_{t,k} X_{t-k}, \quad t = 1, \ldots, N-1,$$

is such that the mean squared prediction error $\sigma_t^2 \equiv E\{(X_t - \widehat{X}_t)^2\}$ is minimized over all possible linear combinations of $X_{t-1}, \ldots, X_0$. Define an $N \times N$ lower triangular matrix L^{-1} whose first row (indexed by $k = 0$) is given by $[1, 0, \ldots, 0]$, while the remaining rows $k = 1, \ldots, N-1$ are given by

$$[-\phi_{k,k}, -\phi_{k,k-1}, \ldots, -\phi_{k,1}, 1, \underbrace{0, \ldots, 0}_{N-1-k \text{ zeros}}].$$

The matrix L^{-1} is taken to be the inverse of the matrix L, which itself must also be lower triangular. Note that multiplication of the tth row of L^{-1} times $\mathbf{X}$ yields $X_t - \widehat{X}_t$, i.e, the prediction error (when $t = 0$, we define $\widehat{X}_0 \equiv 0$). We can express the mean squared prediction error σ_t^2 as $\sigma_\epsilon^2 \gamma_t^2$, where

$$\gamma_t^2 \equiv \frac{\Gamma(1-2\delta)}{\Gamma^2(1-\delta)} \prod_{n=1}^{t} (1 - \phi_{n,n}^2) \quad \text{for } t = 0, \ldots, N-1$$

(when $t = 0$, we interpret the product in the above to be unity). Define D to be an $N \times N$ diagonal matrix whose diagonal elements are $\gamma_0^2, \gamma_1^2, \ldots, \gamma_{N-1}^2$. The vector of prediction errors $L^{-1}\mathbf{X}$ is multivariate Gaussian with a mean vector of zeros and with a covariance matrix given by $\sigma_\epsilon^2 D$; i.e., the prediction errors are pairwise uncorrelated, but are not identically distributed since they have different variances $\sigma_\epsilon^2 \gamma_t^2$. If we let $\boldsymbol{\epsilon}_N \equiv [\epsilon_0, \epsilon_1, \ldots, \epsilon_{N-1}]^T$ be an N dimensional vector containing N independent and identically distributed Gaussian RVs with zero mean and unit variance and if we let $D^{1/2}$ denote the matrix whose elements are the square roots of the corresponding elements of D so that $D^{1/2} D^{1/2} = D$, then $\sigma_\epsilon D^{1/2} \boldsymbol{\epsilon}_N \stackrel{\mathrm{d}}{=} L^{-1}\mathbf{X}$, and hence $\sigma_\epsilon L D^{1/2} \boldsymbol{\epsilon}_N \stackrel{\mathrm{d}}{=} \mathbf{X}$, where '$\stackrel{\mathrm{d}}{=}$' stands for 'has the same distribution as.' We can thus use $\sigma_\epsilon L D^{1/2} \boldsymbol{\epsilon}_N$ to generate a portion of a realization from a stationary FD process using just N RVs from a white noise process despite the fact that the representation for X_t in Equation (11) formally involves an infinite number of white noise RVs.

Let $L_{t,k}$ denote the (t, k)th element of L. Because L is lower triangular, the element $L_{t,k}$

must be zero outside of the choice of indices $t = 0, \ldots, N-1$ and $k = 0, \ldots, t$. We can thus write

$$X_t \stackrel{\mathrm{d}}{=} \sigma_\epsilon \sum_{k=0}^{t} b_{t,k}(\delta)\epsilon_k, \quad t = 0, 1, \ldots, N-1, \tag{14}$$

where

$$b_{t,k}(\delta) \equiv L_{t,k}\gamma_k \text{ for } \delta < 1/2. \tag{15}$$

Note that the above depends on the tth row of L, say, $\mathbf{r}_t^T$. We can obtain this row by solving the equation $L^{-T}\mathbf{r}_t = \mathbf{i}_t$, where L^{-T} denotes the transpose of L^{-1}, and $\mathbf{i}_t$ is a vector, all of whose elements are zero except for the tth element, whose value is one. Since L^{-T} is upper triangular, we can efficiently solve for $\mathbf{r}_t$ using a back substitution scheme (see, e.g., §2.3 of Press *et al.*, 1992).

Let us now turn to the case of a nonstationary FD process $\{X_t\}$ with parameters $1/2 \le \delta < 3/2$ and σ_ϵ^2. By definition, the first difference $X_t^{(1)} = X_t - X_{t-1}$ of this process yields a stationary FD process with parameters $\delta - 1$ and σ_ϵ^2. We can thus simulate a portion of $\{X_t\}$ by cumulatively summing $\{X_t^{(1)}\}$ as per Equation (9):

$$X_t = \sum_{l=0}^{t} X_l^{(1)}, \quad t = 0, 1, \ldots, N-1.$$

Since $X_t^{(1)}$ has the same distribution as the RV on the right-hand side of Equation (14) after we replace δ with $\delta - 1$, we obtain

$$X_t \stackrel{\mathrm{d}}{=} \sigma_\epsilon \sum_{l=0}^{t}\sum_{k=0}^{l} b_{l,k}(\delta-1)\epsilon_k \equiv \sigma_\epsilon \sum_{k=0}^{t} b_{t,k}(\delta)\epsilon_k, \tag{16}$$

where

$$b_{t,k}(\delta) \equiv \sum_{j=k}^{t} b_{j,k}(\delta-1) \text{ when } 1/2 \le \delta < 3/2. \tag{17}$$

The important fact to note is that Equation (14) for the stationary case and Equation (16) for the case of a nonstationary FD process with stationary first differences both involve a finite linear combination of a sequence $\epsilon_0, \ldots, \epsilon_t$ of independent standard Gaussian RVs.

We can generalize this scheme to handle nonstationary FD processes with $\delta \geq 3/2$. Thus, if X_t is an FD process with parameters $3/2 \leq \delta < 5/2$ and σ_ϵ^2, we can write

$$X_t = \sum_{l=0}^{t} X_l^{(1)} \stackrel{\mathrm{d}}{=} \sigma_\epsilon \sum_{k=0}^{t} b_{t,k}(\delta)\epsilon_k, \text{ where now } X_t^{(1)} = \sum_{l=0}^{t} X_l^{(2)},$$

and

$$b_{t,k}(\delta) \equiv \sum_{j=k}^{t} b_{j,k}(\delta - 1) \quad \text{for } 3/2 \leq \delta < 5/2. \tag{18}$$

Since $1/2 \leq \delta - 1 < 3/2$, the weights in the right-hand side of the above can be computed using Equation (17). Processes with $\delta \geq 5/2$ can be handled in an analogous manner.

We can now define a set of RVs that has the same distribution as a TVFD process $\{X_t(\boldsymbol{\beta}_t)\}$ with parameter vector $\boldsymbol{\beta}_t = [\delta_t, \sigma_{\epsilon,t}^2]^T$ such that $\delta_t < 5/2$ (this range for δ_t will cover most cases of practical interest, but we can obviously extend the definition to expand the upper range of δ_t if so desired). We have

$$X_t(\boldsymbol{\beta}_t) \stackrel{\mathrm{d}}{=} \sigma_{\epsilon,t} \sum_{k=0}^{t} b_{t,k}(\delta_t)\epsilon_k, \tag{19}$$

where $b_{t,k}(\delta_t)$ is defined as in Equations (15), (17) or (18) when, respectively, $\delta_t < 1/2$, $1/2 \leq \delta_t < 3/2$ or $3/2 \leq \delta_t < 5/2$.

3.2 Circulant Embedding Approach

Let us now review the circulant embedding scheme for generating a realization from a portion $\mathbf{X}$ of a stationary FD process with parameters $-1 \leq \delta < 1/2$ and $\sigma_\epsilon^2 > 0$ (Davies and Harte, 1987; Wood and Chan, 1994; Dietrich and Newsam, 1997; Chilès and Delfiner, 1999). First, we use Equation (6) with σ_ϵ^2 replaced by unity and then Equation (7) to generate an ACVS $\{s_{X,\tau}\}$ out to lag $\tau = N$. Second, we compute the real-valued sequence

$$S_k \equiv \sum_{\tau=0}^{N} s_{X,\tau} e^{-i2\pi f_k \tau} + \sum_{\tau=N+1}^{2N-1} s_{X,2N-\tau} e^{-i2\pi f_k \tau}, \quad k = 0, 1, \ldots, N, \tag{20}$$

where $f_k \equiv k/(2N)$. Note that the above can be obtained from the discrete Fourier transform (DFT) of the following sequence of length $2N$:

$$s_{X,0}, s_{X,1}, \ldots, s_{X,N-2}, s_{X,N-1}, s_{X,N}, s_{X,N-1}, s_{X,N-2}, \ldots, s_{X,1}.$$

In order for the circulant embedding approach to work, it is required that S_k be nonnegative for $k = 0, 1, \ldots, N$. Results in Gneiting, 1998, and Craigmile, 2002, show that this is true for FD processes such that $-1/2 \leq \delta < 1/2$. Lemma 4.2 in Craigmile, 2002, can easily be extended to handle $-1 \leq \delta < -1/2$ by noting that Equations (6) and (7) hold for all δ. For $\delta < -1$, some of the S_k can be negative. For example, when $N = 2$, we have $S_0 \propto 2(1+\delta)/[(1-\delta)(2-\delta)]$, which is negative for all $\delta < -1$. Third, let $\epsilon_0, \ldots, \epsilon_{2N-1}$ be a set of $2N$ independent standard Gaussian RVs, and construct the following complex-valued sequence:

$$\mathcal{Y}_k \equiv \begin{cases} \sigma_\epsilon \epsilon_0 \sqrt{2NS_0}, & k = 0; \\ \sigma_\epsilon(\epsilon_{2k-1} + i\epsilon_{2k})\sqrt{NS_k}, & 1 \leq k < N; \\ \sigma_\epsilon \epsilon_{2N-1}\sqrt{2NS_N}, & k = N; \\ \mathcal{Y}^*_{2N-k}, & N < k \leq 2N-1; \end{cases}$$

(here the asterisk denotes complex conjugation). Finally, we use the inverse DFT to construct the real-valued sequence

$$Y_t = \frac{1}{2N} \sum_{k=0}^{2N-1} \mathcal{Y}_k e^{i2\pi f_k t}, \quad t = 0, \ldots, 2N-1.$$

We now have

$$[Y_0, Y_1, \ldots, Y_{N-1}]^T \stackrel{\mathrm{d}}{=} \mathbf{X};$$

i.e., we can regard a realization from the first N values of the Y_t sequence to be a realization from a portion of an FD process.

Let $\mathcal{Y}$, $\mathbf{Y}$ and $\boldsymbol{\epsilon}$ be vectors of dimension $2N$ containing, respectively, the $\mathcal{Y}_k$, Y_t and ϵ_k sequences. We can then write $\mathcal{Y} = \sigma_\epsilon G(\delta)\boldsymbol{\epsilon}$, where $G(\delta)$ is a $2N \times 2N$ complex-valued matrix that depends just upon δ through S_k. Let F be a $2N \times 2N$ complex-valued matrix expressing the inverse DFT so that $\mathbf{Y} = F\mathcal{Y}$. Since $\mathbf{Y}$ and $\boldsymbol{\epsilon}$ are real-valued, we have both $\mathbf{Y} = \sigma_\epsilon FG(\delta)\boldsymbol{\epsilon}$ and $\mathbf{Y} = \sigma_\epsilon F^* G^*(\delta)\boldsymbol{\epsilon}$, where F^* is the matrix whose values are the complex conjugates of the corresponding elements in F, and $G^*(\delta)$ is similarly defined. We can thus write

$$\mathbf{Y} = \sigma_\epsilon C(\delta)\boldsymbol{\epsilon}, \text{ where } C(\delta) \equiv \frac{FG(\delta) + F^*G^*(\delta)}{2},$$

and $C(\delta)$ is necessarily real-valued. If we let $c_{t,k}(\delta)$ denote the (t,k)th element of $C(\delta)$, then we have

$$X_t \stackrel{\mathrm{d}}{=} \sigma_\epsilon \sum_{k=0}^{2N-1} c_{t,k}(\delta)\epsilon_k. \tag{21}$$

The equation above is analogous to Equation (14), which is the key to using the modified Cholesky decomposition to simulate stationary FD processes. In order to handle nonstationary FD processes with the circulant embedding approach, we can use the same construction as in the case of the modified Cholesky decomposition; i.e., we need only create a certain number of cumulative sums of an appropriate stationary FD process. With appropriate definitions for $c_{t,k}(\delta)$ when $\delta > 1/2$, this construction expands the range of δ over which Equation (21) holds. The extension to TVFD processes is based upon

$$X_t(\boldsymbol{\beta}_t) \stackrel{\mathrm{d}}{=} \sigma_{\epsilon,t} \sum_{k=0}^{2N-1} c_{t,k}(\delta_t)\epsilon_k, \tag{22}$$

which is the analog of Equation (19).

3.3 'Cut and Paste' Interpretation of Simulation Methods

The simulation of a TVFD process via either the modified Cholesky decomposition or circulant embedding techniques can be accomplished by 'cutting and pasting' simulations from several different FD processes. Let us focus on the circulant embedding approach, although a similar argument can be made for the modified Cholesky decomposition method. To illustrate, we use Equation (22) with $\sigma_{\epsilon,t}$ set to unity to generate a particular realization from a TVFD process, namely,

$$x_t(\delta_t) \equiv \sum_{k=0}^{2N-1} c_{t,k}(\delta_t) e_k, \quad t = 0, \dots, N-1,$$

where the $e_0, \dots, e_{2N-1}$ represent a single realization from a Gaussian white noise process with zero mean and unit variance. Let us assume that δ_t takes on J different values, $\delta^{(j)}$, $j = 0, \dots, J-1$, where necessarily $1 \le J \le N$. Suppose we generate

$$x_t(\delta^{(j)}) \equiv \sum_{k=0}^{2N-1} c_{t,k}(\delta^{(j)}) e_k, \quad t = 0, \dots, N-1,$$

for $j = 0, \dots, J-1$. Then the simulated series $\{x_t(\delta^{(j)})\}$ is a realization from an FD process with parameter $\delta = \delta^{(j)}$. Let $\mathcal{I}_j$, $j = 0, \dots, J-1$, be the set of indices t such that $\delta_t = \delta^{(j)}$ (note that these sets are disjoint and that their union is $\{0, 1, \dots, N-1\}$). To construct $x_t(\delta_t)$, we need only 'cut' the values $x_t(\delta^{(j)}), t \in \mathcal{I}_j$, from the jth simulated FD process and then 'paste' all J sets of such cuts together; i.e., $x_t(\delta^{(j)})$ becomes $x_t(\delta_t)$ for $t \in \mathcal{I}_j$. To handle the more general case where $\sigma_{\epsilon,t}$ is not always unity, we only need multiply $x_t(\delta^{(j)})$ by $\sigma_{\epsilon,t}$ prior to pasting the series together.

4 Examples

Four examples of typical realizations from TVFD processes are shown in Figure 1. Each simulated series is of length $N = 1024$ and was constructed using the circulant embedding method (comparable realizations can be obtained using the modified Cholesky decomposition). For a series of this length, this method requires 2048 independent deviates from a standard Gaussian distribution. All four time series shown in the figure were created using the same set of 2048 deviates.

Each plot in Figure 1 consists of two panels. The upper panel shows how δ_t was varied from $t = 0$ to 1023. In example (a) in the upper left-hand plot, we varied δ_t linearly from 0 up to 0.4 so that it is always contained in the stationary range for an FD process (i.e., $\delta < 0.5$). The plot in the lower panel clearly shows the spectral properties of the process evolving with time. This series is comparable to a realization from a locally stationary process.

In example (b), we set $\delta_t = 0.48$ for the first half of the series and $\delta_t = 0.52$ for the second half. Since the boundary between the stationary and nonstationary FD processes occurs at $\delta = 0.5$, this series starts out as stationary but then transitions into a nonstationary FD process with stationary first differences. Visually, there is little difference in the properties of the process before and after the mid-point of the series. This demonstrates that realizations of short segments from FD processes near the stationary/nonstationary boundary are qualitatively similar. This example was created by simulating one FD process with $\delta = 0.48$ and a second with $\delta = 0.52$ using the same 2048 independent standard Gaussian deviates, after which the first 512 simulated deviates from the first simulation were pasted together with the last 512 simulated deviates from the second simulation. Because of its transition across stationary/nonstationary boundary, this series cannot be considered to be a realization from either a locally stationary or locally self-similar process.

In example (c), we set $\delta_t = 5/6$ for the first half of the series and $\delta_t = 1$ for the second half. The associated power laws are $\alpha = -5/3$ (corresponding to that of Kolmogorov turbulence) and $\alpha = -2$ (corresponding to that of a random walk or Brownian motion process). For plotting purposes, we set $\sigma_\epsilon^2 = 0.16$ here rather than using unity as in the three other examples. This series is an example of a realization from a locally self-similar process.

For the final example (d), we varied δ_t according to a crude piecewise linear approximation to variations observed in a set of aerothermal turbulence measurements (Constantine *et al.*, 2001). Here δ_t starts out in the stationary region, then evolves into a nonstationary region that is intended to mimic Kolmogorov turbulence (i.e., $\delta = 5/6$), after which it returns to the stationary region.

5 Concluding Remarks

We have presented two exact schemes for simulating TVFD processes. Both schemes are well known in the context of simulating stationary FD processes and can be readily extended to simulate nonstationary FD processes via cumulative sums. Here we have shown that the simple device of using the same set of Gaussian white noise deviates to create simulated series with different values of δ and then 'cutting and pasting' the different simulations together yields exact simulations of TVFD processes.

Although use of either scheme yields exact simulations of a TVFD process and thus eliminates one source of potential error in Monte Carlo studies of, e.g., estimators of δ_t, the two methods do have some strengths and weaknesses. First, while the scheme based upon the modified Cholseky decomposition can handle any δ, the one based upon circulant embedding requires $\delta \geq -1$. In our experience, this is not a serious limitation for the circulant embedding method: the case $\delta < -1$ corresponds to time series that are highly antipersistent, which rarely arise in real-world processes. Second, if we ignore the effort required to obtain the weights $b_{t,k}(\delta_t)$ and $c_{t,k}(\delta_t)$ (usually negligible when conducting a

Monte Carlo study involving many simulated series from the same TVFD process), then the computational burden of the modified Cholesky decomposition method is $O(N^2)$, whereas the burden for the circulant embedding method is $O(J \cdot N \cdot \log(N))$, where J is the number of distinct values assumed by δ_t. If J is small compared to N, then the circulant embedding method is much faster to use; on the other hand, if $J = N$ as in example (a) shown in Figure 1, than the method based upon the modified Cholesky decomposition is faster to use. Both methods are thus worth implementing.

Finally, we note that the discussion in §3 can be adapted to define many other classes of processes with time-varying properties. For example, the right-hand side of Equation (11) can be generalized easily to other parameterized weighted summations of white noise (e.g., autoregressive processes). If the parameters are allowed to vary over time, then we can create exact simulations via the modified Cholesky decomposition approach (which will always work) or the circulant embedding method (which will work subject to verification that $S_k \geq 0$ in Equation (20)).

6 Acknowledgments

This material is based upon work supported by the Air Force Office of Scientific Research under an STTR grant to the Insightful Corporation and the University of Washington (contract number No. F49620–99–C–0068). We would also like to thank Prof. Peter Craigmile, Department of Statistics, Ohio State University, and Prof. Tilmann Gneiting, Department of Statistics, University of Washington, for their helpful comments on the manuscript.

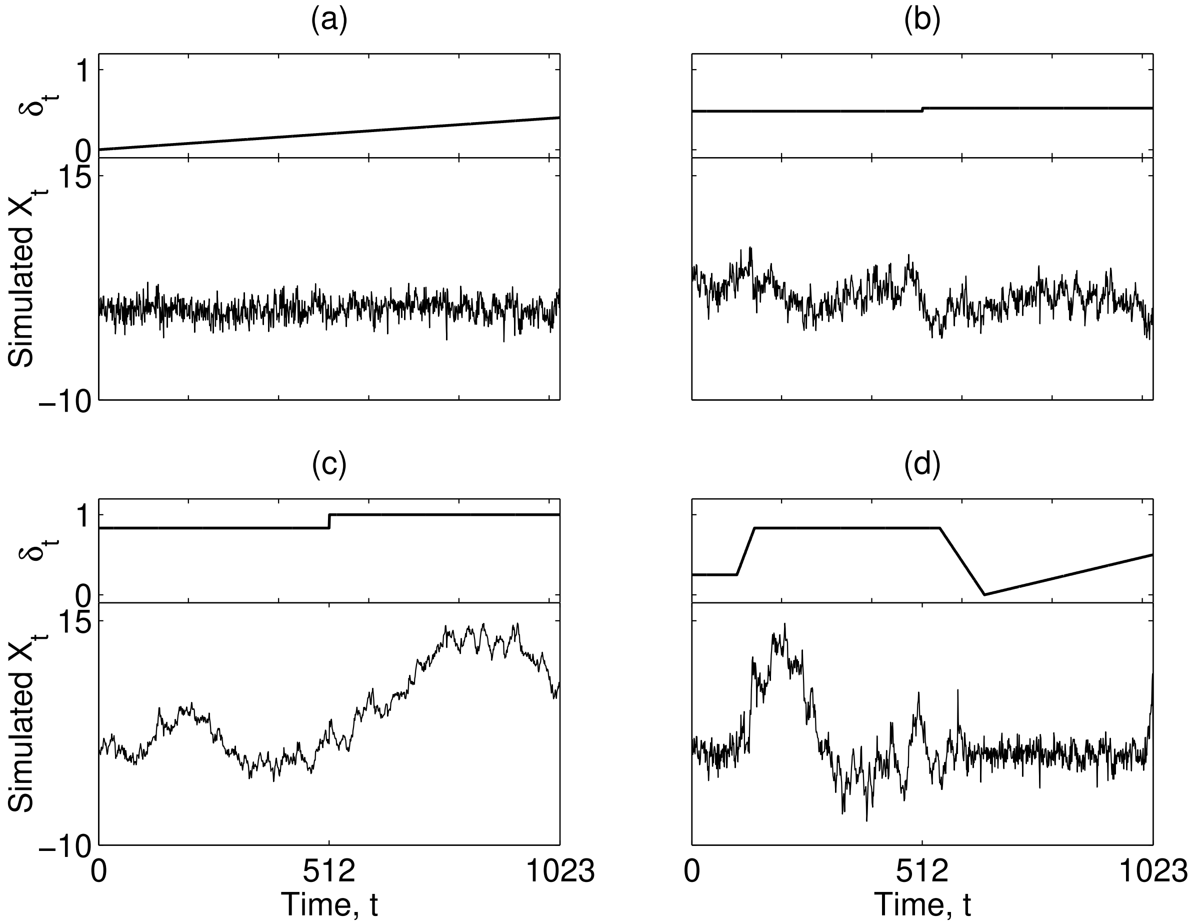

Figure 1: Four examples of time series simulated from TVFD processes. The upper panel in each plot shows the sequence $\delta_t, t = 0, \ldots, 1023$, used to generate the simulated time series in the lower panel. All four simulated series were created using the circulant embedding method on the same set of 2048 independent deviates from a standard Gaussian distribution.

Appendix

Granger and Joyeux, 1980, and Hosking, 1981, prove some basic results about the the ACVS and PACS for stationary FD processes under the restriction that $-1 < \delta < 1/2$. Here we give a theorem and a remark that extend these results to all stationary FD processes (i.e., $\delta < 1/2$).

Lemma 1: If $\{X_t\}$ is any FD process with $\delta < 1/2$ such that $\delta \neq 0, -1, -2, \ldots$, then its ACVS is given by Equation (4).

Proof: The desired result follows if we can show that, for any nonnegative integer d, Equation (4) is true when $\delta \in [-1/2-d, 1/2-d)$, $\delta \neq -d$. Granger and Joyeux, 1980, established that this equation is true for $-1 < \delta < 1/2$, $\delta \neq 0$, so the result holds for $d = 0$. We claim that, if the result is true for some $d \geq 0$, then it must be true for $d+1$ also, which would establish the lemma by induction. To see that this claim is true, suppose $\{Y_t\}$ is an FD process with $-1/2 - d \leq \delta' < 1/2 - d$, $\delta' \neq -d$, and let $X_t \equiv Y_t - Y_{t-1}$. Then $\{X_t\}$ is an FD process with $\delta \equiv \delta' - 1$. i.e., $-1/2 - (d+1) \leq \delta < 1/2 - (d+1)$, $\delta \neq -(d+1)$. The ACVS $\{s_{X,\tau}\}$ of $\{X_t\}$ is related to the ACVS $\{s_{Y,\tau}\}$ of $\{Y_t\}$ as follows:

$$s_{X,\tau} = \operatorname{cov}\{X_t, X_{t-\tau}\} = \operatorname{cov}\{Y_t - Y_{t-1}, Y_{t-\tau} - Y_{t-\tau-1}\} = 2s_{Y,\tau} - s_{Y,\tau+1} - s_{Y,\tau-1}.$$

Using the induction hypothesis, Equation (4) and $\delta' = \delta + 1$, we have

$$s_{X,\tau} = -\frac{\sigma_\epsilon^2 \sin(\pi\delta)\Gamma(-1-2\delta)}{\pi}\left(\frac{2\Gamma(\tau+\delta+1)}{\Gamma(\tau-\delta)} - \frac{\Gamma(\tau+2+\delta)}{\Gamma(\tau+1-\delta)} - \frac{\Gamma(\tau+\delta)}{\Gamma(\tau-1-\delta)}\right).$$

Use of the relationship $\Gamma(x+1) = x\Gamma(x)$ along with some algebra shows that the above reduces to Equation (4), thus establishing the lemma.

Lemma 2: If $\{X_t\}$ is an FD process with parameters $\delta = -d$ and $\sigma_\epsilon^2 > 0$, where d is any nonnegative integer, then the nonzero portion of its ACVS $\{s_{X,\tau}\}$ is given by Equation (5).

Proof: When $\delta = 0$, an FD process reduces to a white noise process $\{\sigma_\epsilon \epsilon_t\}$ with mean zero and variance σ_ϵ^2, for which $s_{X,0} = \sigma_\epsilon^2$ and $s_{X,\tau} = 0$ for $\tau \neq 0$, so the lemma holds in this case. If $\delta = -d$ for some positive integer d, the FD process is the dth order backward difference of $\{\sigma_\epsilon \epsilon_t\}$:

$$X_t = \sigma_\epsilon (1 - B)^d \epsilon_t = \sigma_\epsilon \sum_{k=0}^{d} \binom{d}{k} (-1)^k \epsilon_{t-k}.$$

For $\tau > d$, we have $s_{X,\tau} = 0$; on the other hand, when $0 \leq \tau \leq d$, the ACVS is given by

$$s_{X,\tau} = \sigma_\epsilon^2 \sum_{j=0}^{d} \sum_{k=0}^{d} \binom{d}{j} \binom{d}{k} (-1)^j (-1)^k E\{\epsilon_{t-j} \epsilon_{t+\tau-k}\} = \sigma_\epsilon^2 (-1)^\tau \sum_{j=0}^{d-\tau} \binom{d}{j} \binom{d}{\tau + j}.$$

Equation 2 of §0.156, Gradshteyn and Ryzhik, 1980, shows that the right-hand side reduces to Equation (5), as required.

Theorem: If $\{X_t\}$ is an FD process with $\delta < 1/2$, then its ACVS $\{s_{X,\tau}\}$ is given by Equations (6) and (7).

Proof: Suppose first that δ is not an integer. Then Equation (4) holds, which reduces to Equation (6) when $\tau = 0$. Use of Equation (4) and the relationship $\Gamma(x+1) = x\Gamma(x)$ readily give Equation (7).

Now suppose that $\delta = -d$, where d is a nonnegative integer. Then Equation (5) holds. Letting $\tau = 0$ yields

$$s_{X,0} = \sigma_\epsilon^2 \binom{2d}{d} = \sigma_\epsilon^2 \frac{\Gamma(1+2d)}{\Gamma^2(1+d)} = \sigma_\epsilon^2 \frac{\Gamma(1-2\delta)}{\Gamma^2(1-\delta)},$$

which is in agreement with Equation (6). Since $s_{X,\tau} \neq 0$ when $1 \leq \tau \leq d$, we can write

$$\frac{s_{X,\tau}}{s_{X,\tau-1}} = -\frac{(d-\tau+1)!(d+\tau-1)!}{(d-\tau)!(d+\tau)!} = \frac{\tau - d - 1}{\tau + d} = \frac{\tau - \delta - 1}{\tau - \delta},$$

which is in agreement with Equation (7). Finally, when $\tau = d + 1$, this equation yields

$$s_{X,d+1} = s_{X,d} \frac{d + 1 + \delta - 1}{\tau - \delta} = 0.$$

When $\tau > d+1$, use of Equation (7) correctly yields $s_{X,\tau} = 0$, as required. This establishes the theorem.

Remark: When $\{X_t\}$ is an FD process with $-1/2 < \delta < 1/2$, Hosking, 1981, established that its PACS $\{\phi_{t,t}\}$ is given by Equation (8) using a proof by induction. Based upon the fact that Equations (4) and (5) give the ACVSs for all stationary FD processes, it is possible to use his proof to establish Equation (8) for all $\delta < 1/2$,

References

Beran, Jan (1994). *Statistics for Long-Memory Processes*. Chapman & Hall, New York.

Carmona, R., W.-L. Hwang, and B. Torrésani (1998). *Practical Time-Frequency Analysis*. Academic Press, San Diego.

Chilès, J. and P. Delfiner (1999). *Geostatistics: Modeling Spatial Uncertainty*. Wiley, New York.

Constantine, W. L. B., D. B. Percival, and P. G. Reinhall (2001). "Inertial range determination for aerothermal turbulence using fractionally differenced processes and wavelets". *Physical Review E*, 64(036301) 12 pages.

Craigmile, P. F. (2002). "Simulating a class of stationary Gaussian processes using the Davies–Harte algorithm, with application to long memory processes". *Journal of Time Series Analysis*. To appear.

Dahlhaus, R. (1996). "Maximum likelihood estimation and model selection for nonstationary processes". *Nonparametric Statistics*, 6 171–91.

—— (1997). "Fitting time series models to nonstationary processes". *The Annals of Statistics*, 25 1–37.

Davies, R. B. and D. S. Harte (1987). "Tests for the Hurst effect". *Biometrika*, 74 95–102.

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 Computing*, 18 1088–107.

Flandrin, P. and P. Gonçalvès (1994). "From wavelets to time-energy distributions". In Schumaker, L. L. and G. Webb, eds., *Recent Advances in Wavelet Analysis*, pp. 309–34. Academic Press, Boston.

Flandrin, P. (1989). "On the spectrum of fractional Brownian motions". *IEEE Transactions on Information Theory*, 35 197–9.

Gneiting, T. (1998). "Simple tests for the validity of correlation function models on the circle". *Statistics and Probability Letters*, 39 119–22.

Gonçalvès, P. and P. Flandrin (1993). "Bilinear time-scale analysis applied to local scaling exponent estimation". In Meyer, Y. and S. Roques, eds., *Progress in Wavelet Analysis and Applications*, pp. 271–6. Frontieres, Paris.

Gradshteyn, I. S. and I. M. Ryzhik (1980). *Table of Integrals, Series, and Products (Corrected and Enlarged Edition)*. Academic Press, New York.

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. M. R. (1981). "Fractional differencing". *Biometrika*, 68 165–76.

—— (1984). "Modeling persistence in hydrological time series using fractional differencing". *Water Resources Research*, 20 1898–908.

Mallat, S. G., G. Papanicolaou, and Z. Zhang (1998). "Adaptive covariance estimation of locally stationary processes". *The Annals of Statistics*, 26 1–47.

Mallat, S. G. (1999). *A Wavelet Tour of Signal Processing (Second Edition)*. Academic Press, San Diego.

Mandelbrot, B. B. and J. W. van Ness (1968). "Fractional Brownian motions, fractional noises and applications". *SIAM Review*, 10(4) 422–37.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical Recipes in Fortran (Second Edition)*. Cambridge University Press, Cambridge, UK.

Priestley, M. B. (1965). "Evolutionary spectra and non-stationary processes". *Journal of the Royal Statistical Society, Series B*, 27 204–37.

Sinai, Y. G. (1976). "Self-similar probability distributions". *Theory of Probability and Its Applications*, 21 64–80.

Wang, Y., J. E. Cavanaugh, and C. Song (2001). "Self-similarity index estimation via wavelets for locally self-similar processes". *Journal of Statistical Planning and Inference*, 99 91–110.

Whitcher, B. J and M. J. Jensen (2000). "Wavelet estimation of a local long memory parameter". *Exploration Geophysics*, 31 94–103.

Wood, A. T. A. and G. Chan (1994). "Simulation of stationary Gaussian processes in $[0,1]^d$". *Journal of Computational and Graphical Statistics*, 3 409–32.

Yaglom, A. M. (1958). "Correlation theory of processes with random stationary nth increments". *American Mathematical Society Translations, Series 2*, 8 87–141.