Nonlinear Processes in Geophysics (1998) 5: 63–68

Nonlinear Processes in Geophysics

Bounded lognormal cascades as quasi-multiaffine random processes

M. Menabde

Department of Physics, University of Auckland, Private Bag 92019, Auckland, New Zealand

Received: 13 June 1997 – Revised: 16 May 1998 – Accepted: 14 September 1998

Abstract. A new model for simulating non-negative random processes is developed. The model is based on a bounded random multiplicative cascade with a lognormal generator with the scale parameter dependent on the number of steps in the cascade. It is shown, both analytically and numerically, that the simulated field has multiaffine properties within some restricted range of scales. This type of random processes is defined here as quasi-multiaffine. A procedure for retrieving the modeling parameters from real data is also discussed.

1 Introduction

Multiscaling models have been increasingly popular in different areas of geophysics in the last decade. They have been used for the statistical description and simulation of atmospheric turbulence, rainfall processes, water distribution in clouds and in many other applications. These models are based on the empirical evidence about scaling properties of many geophysical fields and the analogy with the random multiplicative cascade models in fully developed turbulence. Cascade models are phenomenological, though there have been some attempts to justify them by a quite remote analogy with the energy cascade process in fully developed turbulence, the latter having a semi-phenomenological character itself. For most geophysical applications the existence of a cascade-type process in space and/or time and its physical origin is even less clear. In our view, at present stage, the best possible justification for the use of cascade-type models should be based on the following two properties:

- the ability to provide a robust statistical description of random field incorporating its scaling properties with few parameters, and,
- the ability to simulate a synthetic random field that reproduces the observed statistical properties based on the derived parameters.

We shall consider here a non-negative one-dimensional random field, $F(x)$, where x can represent the time or space coordinate of a one-dimensional field or a one-dimensional cut through a multi-dimensional field (e.g., rainfall time series, liquid water content in clouds, temperature fields etc.). A primary means of identifying the scaling properties of a random field is its power spectrum, $P(k)$, which must have a scaling behaviour:

$$P(k) \propto k^{-\beta}. \tag{1}$$

As we will see below, the power spectrum exponent β plays an important role in the classification of scaling random fields.

The more sophisticated multiscaling characteristics of a random field are:

- the continuous spectra of exponents $K(p)$ defined as

$$\left\langle F_l^p(x) \right\rangle \propto l^{-K(p)}, \tag{2}$$

where $F_l(x)$ denotes the field average over the interval l centered at x, and <...> denotes the ensemble average;

- the continuous spectra of exponents $\zeta(p)$ defined as

$$G_p(l) = \left\langle \left| F(x+l) - F(x) \right|^p \right\rangle \propto l^{\zeta(p)}. \tag{3}$$

The left-hand side of the equation (3) is usually referred to as a generalized structure function. Following Benzi et al. (1993) we shall refer to the random fields showing a scaling behaviour in the sense (3) as multiaffine. It is important to note that (2) and (3) define two different and mutually exclusive types of scaling fields. In the definition (2) $F(x)$ is a positively defined stationary random field with a power spectrum exponent $\beta < 1$ (e.g., Menabde et al., 1997a), whereas (3) cannot be realized for nonnegative and stationary random fields for all l-s (e.g., Veneziano et al., 1996). Cascade models, presented in the literature, such as discrete cascades (e.g., Gupta and Waymire, 1993) and the continuous cascades of Shertzer and Lovejoy (1987) produce multiscaling random fields satisfying (2) with the power spectrum exponent $\beta < 1$. Many geophysical fields, however, usually exhibit $\beta > 1$. Several authors (e.g., Marshak et al., 1994; Davis et al., 1996) argued that the power spectrum exponent in the typically observed range $1 < \beta < 3$ can be used as an indicator of nonstationarity of the random field and its statistics can be described by (3). Shertzer and Lovejoy (1987) used power-law filtering in the Fourier space in order to produce fields with a required β. This procedure is, however, ambiguous as the zero frequency component of the field remains undefined and negative values are produced. The field, of course, can be made nonnegative by inverting the negative values or redefining the zero-frequency component, but the scaling properties of the field will then change and they will not be described by the original parameters, used in simulation.

On the other hand, we have recently demonstrated (Menabde et al., 1997b) the existence of a positively defined stationary random process with an asymptotically scaling power spectrum in some range, $k >> k_0$ and with a power exponent, $\beta > 1$. Such a process, being a stationary random process, may have only an asymptotically multiscaling structure function, as in (3), in some restricted range $l << 1/k_0$. This condition does not seem to be very restrictive, since in practice real geophysical fields only exhibit scaling, if any, within a limited range of scales. It is the aim of this paper to introduce a new model, based on the multiplicative bounded random cascade, which produces a positively defined stationary random process, yet having a multiscaling structure function $G_p(l)$ in some restricted interval. We propose to refer to such type of processes as quasi-multiaffine random processes.

2 Theory

Consider a homogeneous distribution of a field, F_0, over an interval L. On the first step of the cascade we divide the initial interval into two halves (the theory can be easily generalized to the case of an arbitrary branching number) and assign to each of them value $F_1 = F_0 w(1)$ and $F_2 = F_0 w(2)$, where $w(j)$ are different realizations of a random variable W produced by some generator with a probability density, $g(W)$, and having $\langle W \rangle = 1$. On the next step each half is again divided into halves and the procedure is repeated N times leading to a discrete random field:

$$F_N(j_1, j_2,..,j_N) = F_0 w_1(j_1) w_2(j_1, j_2)\ldots w_N(j_1, j_2, \ldots, j_N) \qquad (4)$$

where F_N represents the piecewise constant function on the N intervals of length $l_N = 2^{-N}L$ and the set of binary indexes $j_1, j_2,\ldots, j_N$ indicate 2^N different possible realizations of the random field. In contrast to the commonly used self-similar cascade models, we will use a generator, $g(W)$, which explicitly depends on the number of steps in the cascade and has the scale parameter decreasing with increasing resolution. This type of cascade is called a bounded random cascade (Cahalan et al., 1990; Marshak et al., 1994; Menabde et al., 1997b). Different types of generators can be used to build a model, and due to its phenomenological character, there are no fundamental theoretical reasons for preferring any of them. From a practical point of view we need a generator which, on one hand, has enough adjustable free parameters and on the other hand, is simple enough to be treated analytically. We will consider below a simple case of a lognormal generator. On the n^{th} step of the cascade the generator has the form:

$$W_n = C_n \exp(-\sigma_n X), \qquad (5)$$

where C_n is a normalization constant, defined from the condition

$$\langle W_n \rangle = 1, \qquad (6)$$

σ_n is a scale parameter decreasing with n, and $X = Norm(1, 0)$, is a normal random variable with a unity width and a zero shift. The moments of the lognormal distribution are given by (e.g. Devroye, 1996):

$$\langle \exp(tX) \rangle = \exp(t^2). \qquad (7)$$

Conditioned by (6) the generator (5) takes the form:

$$W_n = \exp(-\sigma_n^2 - \sigma_n X). \qquad (8)$$

We choose the scale parameter σ_n in (5) to be exponentially decreasing with the cascade step:

$$\sigma_n = \sigma 2^{-nH}, \qquad (9)$$

and, therefore, our model has two free parameters, σ and H. Normalization of the field to the required mean intensity provides the third free parameter for the model.

The normal random variables have an important property - stability under addition:

$$\sum_{i=1}^{n} \sigma_i X_i = \sigma_{1;n} X, \qquad (10)$$

where X_i are independent identically distributed copies of X, and

$$\sigma_{1;n}^2 = \sum_{i=1}^{n} \sigma_i^2. \qquad (11)$$

Using the property (10) we can derive the one-point statistics of the simulated field. From (8)–(11) we get that the field values at the highest resolution in the limit of large N are distributed as a lognormal random variable $\exp(-\sigma_0^2 - \sigma_0 X)$ with the scale parameter σ_0 given by the formula:

$$\sigma_0^2 = \sum_{i=1}^{\infty} \sigma_i^2 = \frac{\sigma^2}{4^H - 1}. \qquad (12)$$

The properties of the two-point statistics can be derived analytically under some assumptions specified below. Considering all pairs of points separated by a distance $l_n = L\, 2^{-n}$, we get the following expression for the structure function:

$$G_p(l_n) = \sum_{k=1}^{n} A_k^n G_p^k, \qquad (13)$$

where the combinatorial factor A_k^n is given by

$$A_k^n = \frac{2^{k-1}}{2^n - 1}, \qquad (14)$$

and

$$G_p^k = \lim_{N\to\infty}\left\langle W_1^p..W_{k-1}^p \left|W_k..W_N - W_k'..W_N'\right|^p\right\rangle =$$

$$= \left\langle W_{1;k-1}^p\right\rangle\left\langle\left|W_{k;\infty} - W_{k;\infty}'\right|\right\rangle \qquad (15)$$

Here $W_{1;k-1}$ and $W_{k;\infty}$ are given

$$W_{1;k-1} = \exp(-\sigma_{1;k-1}^2 - \sigma_{1;k-1}X), \qquad (16)$$

and,

$$W_{k;\infty} = \exp(-\sigma_{k;\infty}^2 - \sigma_{k;\infty}X), \qquad (17)$$

where the scale factors $\sigma_{1;k-1}$ and $\sigma_{k;\infty}$ are defined as in (10).

Consider now the case of small distances l_n, i.e. for n satisfying the condition:

$$n >> 1/H. \qquad (18)$$

In this case, as it follows from (13) and (14), the asymptotically largest contribution to the structure function comes from the terms with k close to n. The first factor in the formula (15) can approximately be calculated as:

$$\left\langle W_{1;k-1}^p\right\rangle = \exp\{\frac{\sigma^2(p^2-p)}{4^H-1}\}. \qquad (19)$$

The second factor in the formula (15) can be calculated if we notice that for n satisfying (18), the lognormal random variables $W_{k;\infty}$ can be replaced by the uniform random variables, Y_k, with the probability distribution function

$$p(y) = \begin{cases}1/2\varepsilon_k, & |1-y| \le \varepsilon_k \\ 0, & |1-y| > \varepsilon_k\end{cases}. \qquad (20)$$

Indeed, when the condition (18) holds and for small enough p, moments of the lognormal random variable are approximately equal to

$$\left\langle W_{k;\infty}^p\right\rangle \approx 1 + \sigma_{k;\infty}^2(p^2-p). \qquad (21)$$

On the other hand, in the same approximation the moments of the uniform distribution are given by:

$$\left\langle Y_k^p\right\rangle = \frac{1}{2\varepsilon_k}\int_{1-\varepsilon_k}^{1+\varepsilon_k} y^p dy \approx 1 + \frac{1}{6}\varepsilon_k^p(p^2-p) \qquad (22)$$

Comparing formulae (21) and (22) we can see that we have to take

$$\varepsilon_k = \sqrt{6}\sigma_{k;\infty} = \frac{\sigma\sqrt{6}}{\sqrt{1-4^{-H}}}2^{-kH}. \qquad (23)$$

Now we can calculate the second factor in the formula (15). The difference of two uniform random variables Y_1 and Y_2 with a width ε_k is a random variable $\Delta Y = Y_1 - Y_2$, with a triangular probability density function (Springer, 1979):

$$p(\Delta y) = \begin{cases}\frac{1}{2\varepsilon_k}(1 - |\Delta y|/2\varepsilon_k), & |\Delta y| < 2\varepsilon_k; \\ 0, & |\Delta y| \ge 2\varepsilon_k\end{cases}. \qquad (24)$$

In this approximation for W we get:

$$\left\langle\left|W_{k;\infty} - W_{k;\infty}'\right|^p\right\rangle = \left\langle|\Delta Y|^p\right\rangle = \frac{2^p}{(p+1)(p+2)}\varepsilon_k^p \qquad (25)$$

Substituting (14), (19) and (25) into (13) we get the following expression for the structure function:

$$G_p(l_n) = \frac{f(p)}{2^n-1}\frac{1-c^{-n}}{c-1}, \qquad (26)$$

where $c = 2^{-pH-1}$, and $f(p)$ is some function of p, which explicit form is irrelevant to further consideration.

Consider now the ratio

$$\frac{G_p(2l_n)}{G_p(l_n)} = \frac{G_p(l_{n-1})}{G_p(l_n)} = \frac{(2^n-1)(1-c^{n-1})}{(2^{n-1}-1)(1-c^n)}. \qquad (27)$$

In the limit of large n we get:

$$\lim_{n\to\infty}\frac{G_p(2l_n)}{G_p(l_n)} = \begin{cases}2^{pH}, & p < 1/H; \\ 2, & p > 1/H;\end{cases} \qquad (28)$$

Thus, the structure function exponent has the form:

$$\zeta(p) = \begin{cases}pH, & p < 1/H; \\ 1, & p > 1/H;\end{cases} \qquad (29)$$

As it follows from (9) and (21) this results are valid for the distances l_n satisfying the condition

$$\sigma^2(p^2-p)\left(\frac{l_n}{L}\right)^2 << 1. \qquad (30)$$

The same result for the structure function was derived by Marshak et al., (1994) for a different type of a bounded cascade model. This fact probably indicates that the form (29) of a structure function exponent is universal for different types of bounded cascades models, with the width of the generator exponentially decreasing with the number of steps.

The results of some simulations are presented below in the Section 3, and they confirm that the structure function is scaling over a limited range. The power spectrum of this process has an asymptotically scaling behaviour for large frequencies with the power exponent β=1+$\zeta(2)$ (see Appendix). The analytical properties of one- and two-point statistics of $G_p(l)$, as given, respectively, by (12) and (29), provide a method for the model parameter estimation from real data. From the slope of the structure function at the origin we can find the H parameter, and by fitting the probability distribution of a given data set by a lognormal distribution we can find the σ parameter.

3 Simulation

In order to illustrate the theoretical results we have performed some simulations using a 15-step cascade model with a lognormal generator, thus producing series of 32,768 data points. The typical realizations of this field with two different sets of parameters are shown in Figure 1. To

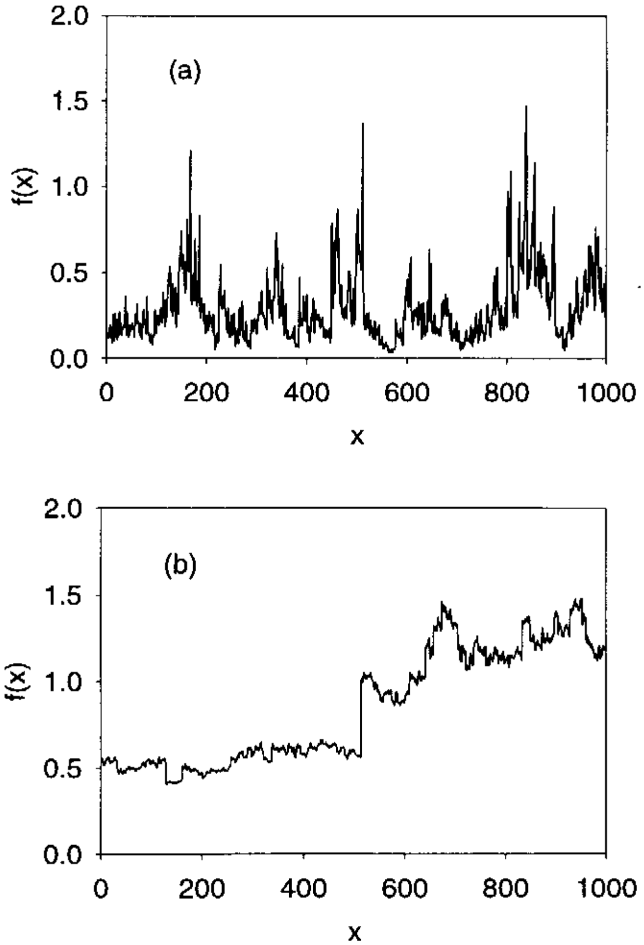

Fig. 1. Two different realizations of the field simulated by the bounded random cascade: (a) $\sigma = 1.0$, $H = 0.2$, (b) $\sigma = 1.0$, $H = 0.5$.

estimate the statistical characteristics for each set of parameters we simulated an ensemble of 100 realizations of the random field for a given set of modeling parameters. The exceedence probability curves for different σ and the same H are shown in Figure 2.

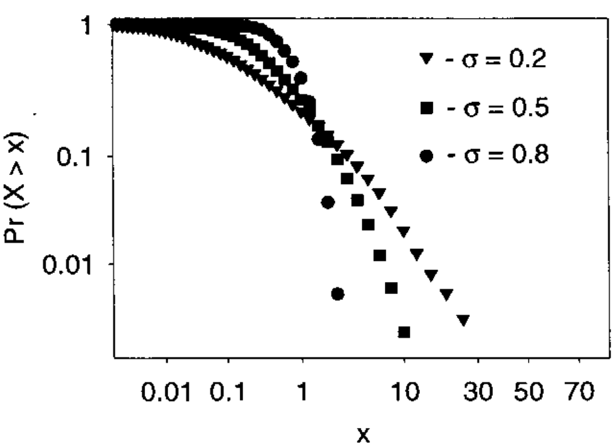

Fig. 2.. Exceedence probability for different realizations of the random field with $H = 0.35$ and different σ.

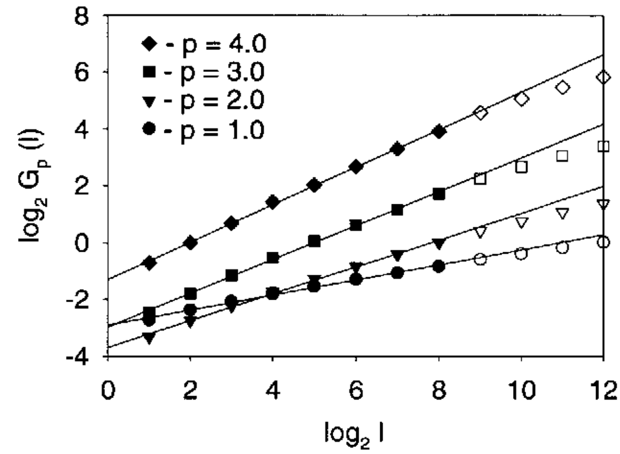

Fig. 3. Scaling of moments of the structure function for the field with $\sigma = 1.0$, $H = 0.35$.

Figure 3 represents a typical behavior of different moments of the structure function. We can see that in this case the scaling range is approximately $0 < l < l_0 2^8$, where l_0 is the smallest scale resolution. Figure 4 shows the structure function behavior for the same σ and different H.

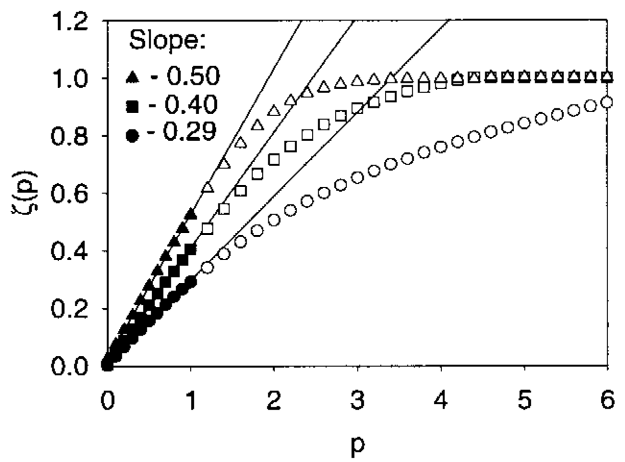

Fig. 4. Structure function for different realization of the random field (●- $\sigma = 1.0$, $H = 0.2$; ■- $\sigma = 1.0$, $H = 0.35$; ▼ $\sigma = 1.0$, $H = 0.50$)

We can see that the slopes of the curves for $H < 0.5$ are biased, as compared with the theoretical prediction (29), due, probably to the finite number of steps in the cascade. However, the dependence of the slope on the modeling parameter H can be established numerically for any given σ, which can be estimated independently from the exceedence probability calculated from the experimental data. Figure 5 shows the power spectra for the same field realizations as in Figure 4. The power spectra $P(k)$ were

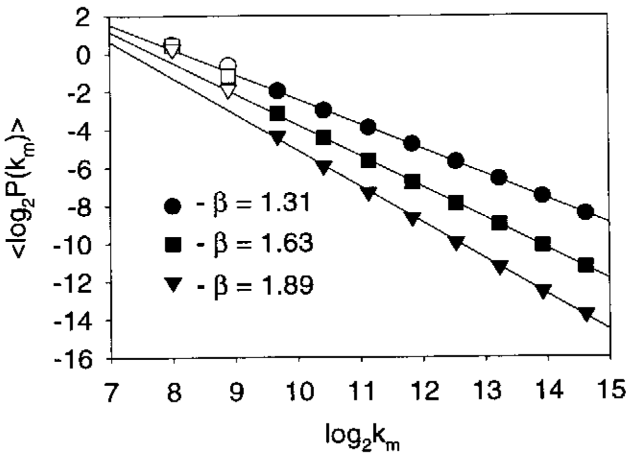

Fig. 5. Power spectra for the same field realizations as in Figure 4.

calculated using the standard FFT algorithm, and then averaged using an octave bining procedure analogous to that found in Davis et al. (1996),

$$\left\langle \log P(\bar{k}_m) \right\rangle = \frac{1}{2^m} \sum_{j=2^m}^{2^{m+1}-1} \log P(k_j)\,, \tag{31}$$

where

$$\log \bar{k}_m = \frac{1}{2^m} \sum_{j=2^m}^{2^{m+1}-1} \log k_j\,. \tag{32}$$

As we can see from Figure 5, the value of β substantially deviates from the theoretical prediction given by (A5), for $H < 0.5$. Moreover, we have found out that β depends on the type of the averaging procedure used while calculating the power spectrum. In our view, the power spectrum should be used as a qualitative, rather than a quantitative characteristic of the random process, with $\beta = 1$ as a threshold value. Power spectrum with $\beta > 1$ indicates that the process can be described by a quasi-multiaffine model, while one with $\beta < 1$ can be described as a self-similar random field (e.g. Menabde et al., 1997a).

Conclusion

We have introduced a new, quite general type of stochastic model based on the bounded random cascade with a lognormal generator. It produces a nonnegative stationary random process which has a multiscaling structure function within some restricted range of scales. We refer to this class of random processes as quasi-multiaffine. This type of behaviour is quite typical for a wide class of geophysical processes and the model has potential as a useful tool for their statistical description and simulation. The model has 3 free parameters and provides a simple procedure for their estimation, based on fitting the one- and two-point statistics of the real data.

Appendix. Power spectrum of quasi-multiaffine random process.

Consider an isotropic and homogeneous 1D random field $F(x)$ which can be represented as a Fourier-Stieltjes integral

$$F(x) = \int \exp(ikx) Z(dk)\,, \tag{A1}$$

where $Z(dk)$ is a random measure satisfying the condition

$$\left\langle Z(dk)\bar{Z}(dq) \right\rangle = P(k)\delta(k-q)dkdq\,. \tag{A2}$$

Here $\bar{Z}$ denotes the complex conjugate, $\delta(k\text{-}q)$ is the delta-function, and $P(k)$ is the power spectrum. Substituting (A1) and (A2) to the formula (3) we get:

$$G_2(l) = \left\langle \left| F(x+l) - F(x) \right|^2 \right\rangle = 8\int_0^{\infty} P(k)\sin^2(kl/2)dk\,. \tag{A3}$$

If $P(k) \propto k^{-\beta}$ for some $k >> k_0$, then for distances $l << 1/k_0$, neglecting the non-scaling terms of order $(k_0 l)^2$, from (A3) we get:

$$G_2(l) \propto l^{\beta-1}\,, \tag{A4}$$

and

$$\beta = 1 + \zeta(2). \tag{A5}$$

Acknowledgements. I thank Daniel Harris, Alan Seed, and Geoff Austin for fruitful discussions.

References

Cahalan, R., Nestler, M., Rigway, W., Wiscombe, W., and Bell, T., Marine stratocumulus spatial structure, in *Proceedings of the 4-th International Meeting on Statistical Climatology,* edited by J. Sansom, pp. 28-32, New Zealand Meteorological Service, Wellington, N.Z., 1990.

Benzi, R., Biferale, L., Crisanti, A., Paladin, G., Vergassola, M., and Vulpiani, A., A random process for the construction of multiaffine field, *Physica D, 65,* 353-358, 1993.

Davis, A., Marshak, A., Wiscombe, W., and Cahalan, R., Scale Invariance of Water Distribution in Marine Stratocumulus. Part I; Spectral Properties and Stationarity Issues, *J. Atmos. Sci., 53,* 1538-1558, 1996.

Devroye, L., *Non-Uniform Random Variate Generation,* Springer, 843 pp., New York, 1986.

Gupta, V.K., and Waymire, E.C., A statistical analysis of mesoscale rainfall as a random cascade, *J. Appl. Meteorol., 32,* 251-267, 1993.

Marshak, A., Davis, A., Cahalan, R., and Wiscombe, W., Bounded cascade models as nonstationary multifractals, *Phys. Rev. E, 49,* 55-69, 1994.

Menabde, M., Seed, A., Harris, D., and Austin, G., Self-Similar random fields and rainfall simulation, *J. Geophys. Res, 102,* D12, 13,509-13,515, 1997a.

Menabde, M., Harris, D., Seed, A., and Austin, G., Multiscaling properties of rainfall and bounded random cascades, *Wat. Resour. Res., 33,* No.12, 2823-2830, 1997b.

Schertzer, D., and Lovejoy, S., Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, *J. Geophys. Res., 92,* 9693-9714, 1987.

Springer, M.D., *The algebra of random variables,* John Wiley & Sons, 470 pp., 1979.

Veneziano, D., Bras, R.L., and Niemann, J.D., Nonlinearity and self-similarity of rainfall in time and stochastic model, *J. Geophys. Res., 101D, 26,* 371-26, 392, 1996.