JP4.16 AN ITERATIVE METHOD FOR GENERATING SCALING LOG-NORMAL SIMULATIONS

Gregory M. Lewis*and Philip H. Austin
Department of Earth and Ocean Sciences, University of British Columbia

1. INTRODUCTION

It has been shown that the single-point probability density functions of the data of certain cloud properties, such as cloud optical depth, are positively skewed and resemble the probability density functions of a log-normal random variable (e.g. Barker et al., 1996). There is also evidence that, for certain spatial scales, the two-point statistics are such that the energy spectrum E_k satisfies a power-law relation, with scale invariant exponent β lying between 1 and 3, i.e. $E_k \propto k^{-\beta}$, $1 < \beta < 3$, where the energy spectrum is the modulus squared of the Fourier transform of the data, and k is the wave number.

A Gaussian random field with any desired energy spectrum may be easily generated using Fourier-space filtering due to the fact that the Gaussian distribution is stable (i.e., the sum of any number of Gaussian random variables is a Gaussian random variable). A log-normal simulation may then be generated by exponentiating this Gaussian random field. However, a log-normal field generated in this manner will not have the same energy spectrum as the original or 'mother' Gaussian field. In fact, in general, the form of the energy spectrum will also be affected (e.g. if the energy spectrum of the Gaussian field satisfies a power-law, then the log-normal field will not).

Ideally, it would be possible to specify a target energy spectrum, and have a method to analytically compute the energy spectrum of a mother Gaussian field that will lead to a log-normal field with this target energy spectrum. This approach is taken by Evans et al. (1999) in the case where an auto-correlation function, which depends only on the distance between the field values, can be specified. Thus, the technique of Evans et al. (1999) is limited to the simulation of stationary fields. In the case of scaling random fields, stationarity can be satisfied by restricting $\beta < 1$, or alternatively, by introducing an integral scale (a "scale break") beyond which the variance does not increase. (see, e.g. Davis et al., 1994; Frisch, 1995).

Below, we outline a technique that can be used to generate non-stationary scaling log-normal discrete random fields with scale invariant exponents in the range $1 < \beta < 3$ (i.e. discrete random fields with log-normal single-point statistics and energy spectra that satisfy $E_k \propto k^{-\beta}$, $1 < \beta < 3$). The random fields generated using this new technique are useful for the simulation of cloud data, when, for instance, the origin of a possible scale break is itself the topic of investigation (see, e.g. Lewis et al., 2002).

We extend the range of validity to $1 < \beta < 3$ by using a simple iterative method to numerically approximate the appropriate energy spectrum for the Gaussian field. The iteration produces a sequence of Gaussian random fields, where the difference between the target energy spectrum (for the log-normal field) and the energy spectrum of the exponentiation of a field in the sequence is used to generate the energy spectrum for the next Gaussian field in the sequence (i.e. the next iterate). The main assumption of the method is that the exponentiation of the Gaussian field does not substantially alter the form of the energy spectrum. This is a valid assumption when the variance of the Gaussian field is not large, which is the case of interest for cloud statistics.

In Section 2, we describe the procedure for generating Gaussian random fields. We introduce the iterative method for simulating scaling log-normal random fields in Section 3, and discuss implementation in Section 4. Here, the variance range for which convergence occurs is also discussed. In Section 5, some results of the method are presented.

2. GAUSSIAN RANDOM FIELDS

Our method for generating scaling log-normal simulations requires the exponentiation of Gaussian random fields. Therefore, in this section, we describe the process of simulating a discrete random field that has identically distributed single-point Gaussian statistics and a pre-specified energy spectrum. Because the Gaussian distribution is stable, the result of filtering a Gaussian random field in Fourier-space is a Gaussian random field. Thus, Fourier-space filtering can be used to generate a Gaussian random field with an energy spectrum of any form.

The steps involved in creating a discrete Gaussian random field $f(x_n) \equiv f_n$ (where x_n, n integer, are the N discrete locations of the field values in real space), with mean μ and variance σ^2, are:

1. Create an uncorrelated Gaussian noise, g_n, with mean μ and variance σ^2.

2. Compute the discrete Fourier transform of g_n to obtain the Fourier coefficients $\hat{g}_k$, where k is the wave number. For example, for a one-dimensional field

*Corresponding author address: Gregory Lewis c/o Philip Austin, Atmospheric Science Programme, Department of Earth and Ocean Sciences, University of British Columbia, Vancouver,B.C., V6T 1Z4, Canada; email: paustin@eos.ubc.ca

with $x_n = \{n \; : \; n = 0, 1, ..., N-1\}$,

$$\hat{g}_k = \sum_{n=0}^{N-1} g_n e^{2\pi i n k/N}, \;\; k = 0, 1, ..., N-1.$$

3. Specify a noise-free function $E_k^{(t)}$ that determines the k-dependence of the desired energy spectrum, and is normalized so that the proper variance σ^2 is maintained (see below).

4. Filter the Fourier transformed noise $\hat{g}_k$ with $\sqrt{E_k^{(t)}}$ by simply multiplying to obtain $\hat{f}_k$. That is, $\hat{f}_k = \hat{g}_k \sqrt{E_k^{(t)}}$.

5. Compute the inverse Fourier transform of $\hat{f}_k$ to produce f_n, the Gaussian random field with the desired energy spectrum, which will be $\langle |\hat{f}_k|^2 \rangle = \langle |\hat{g}_k \sqrt{E_k^{(t)}}|^2 \rangle = \langle |\hat{g}_k|^2 \rangle E_k^{(t)} = N\sigma^2 E_k^{(t)}$, where $\langle |\hat{g}_k|^2 \rangle = N\sigma^2$ is independent of k (because g_n is uncorrelated noise), and the angled braces, $\langle \cdot \rangle$, represent an ensemble average. Thus, the energy spectrum of f_n will, on average, be the specified function $E_k^{(t)}$ up to a multiplicative constant.

If the function $E_k^{(t)}$ is not properly normalized, then the (filtered) Gaussian random field f_n will not have the correct variance σ^2. It can be shown that if the normalization condition

$$\frac{1}{N} \sum_{k=1}^{N-1} E_k^{(t)} = 1. \tag{1}$$

is satisfied, the field f_n will, on average, have the same variance σ^2 as the uncorrelated field g_n. In addition, to ensure that f_n has, on average, the correct mean μ, we must have $E_{k=0}^{(t)} = 1$.

Note that, by specifying $E_k^{(t)}$, we are not actually specifying the energy spectrum of the simulation, but we are only determining an ensemble average energy spectrum (i.e. $\langle |f_k|^2 \rangle = N\sigma^2 E_k^{(t)}$). Thus, only on average will the energy spectrum be $\langle |f_k|^2 \rangle$, and each simulation (or realization) will have an energy spectrum that fluctuates about this average spectrum.

Finally, although the discussion seemed to imply that we are generating one-dimensional fields, the process is identical for fields of any dimension.

3. SCALING LOG-NORMAL RANDOM FIELDS

In this section, we introduce the iterative method for generating identically distributed scaling log-normal discrete random fields. In particular, we address the issue of finding the energy spectrum that a Gaussian field must have, so that exponentiation of this field leads to a scaling log-normal field. That is, we are looking for an ensemble average energy spectrum $\langle |\hat{h}_k|^2 \rangle$ for a 'mother' Gaussian field $h(x_n) \equiv h_n$, such that the 'daughter' log-normal random field

$$H_n = e^{h_n}, \tag{2}$$

has energy spectrum

$$E_k^H = \langle |\hat{H}_k|^2 \rangle = c_H^2 k^{-\beta}, \tag{3}$$

where, c_H and β are independent of k, and, here and below, the hatted (discrete) functions are the discrete Fourier transforms of the corresponding real space functions, e.g. $\hat{h}_k$ is the Fourier transform of h_n.

Generate a mother Gaussian field $f_n^{(0)}$, with energy spectrum $\langle |\hat{f}_k^{(0)}|^2 \rangle$, which is a guess for $\langle |\hat{h}_k|^2 \rangle$. Calculate the daughter log-normal field $F_n^{(0)} = e^{f_n^{(0)}}$, and its corresponding Fourier transform, $\hat{F}_k^{(0)}$.

Then, the difference $\hat{d}_k^{(0)}$ between the target energy spectrum (3) and the energy spectrum $\langle |\hat{F}_k^{(0)}|^2 \rangle$ of the log-normal field is

$$\hat{d}_k^{(0)} = \langle |\hat{H}_k|^2 \rangle - \langle |\hat{F}_k^{(0)}|^2 \rangle. \tag{4}$$

Expanding H_n in a Taylor series, we have $H_n = e^{h_n} = 1 + h_n + \text{ h.o.t.}$, where h.o.t. = higher order terms in h_n. If a similar expansion is assumed for $F_n^{(0)}$, then, due to the linearity of the Fourier transform, we have, to within a constant factor,

$$\langle |\hat{H}_k|^2 \rangle - \langle |\hat{F}_k^{(0)}|^2 \rangle = \langle |\hat{h}_k|^2 \rangle - \langle |\hat{f}_k^{(0)}|^2 \rangle + \text{ h.o.t.} \tag{5}$$

If it is assumed that the nonlinearity introduced by exponentiating the Gaussian random field is small (i.e. if h_n and $f_n^{(0)}$ have small magnitude everywhere), then the higher order terms can be ignored. That is, to a linear approximation, the difference in the energy spectra of the daughter log-normal fields is equal to the difference in the energy spectra of the corresponding mother Gaussian fields. Rearranging the terms in (5), and using (4), we obtain

$$\langle |\hat{h}_k|^2 \rangle = \langle |\hat{f}_k^{(0)}|^2 \rangle + \hat{d}_k^{(0)} + \text{ h.o.t.} \tag{6}$$

This suggests that a possible better approximation for $\langle |\hat{h}_k|^2 \rangle$ is $\langle |\hat{f}_k^{(1)}|^2 \rangle$ which satisfies

$$\langle |\hat{f}_k^{(1)}|^2 \rangle = \langle |\hat{f}_k^{(0)}|^2 \rangle + \hat{d}_k^{(0)}. \tag{7}$$

However, more generally (i.e. if the higher order terms are not negligible), $\hat{d}_k^{(0)}$ can be used to define a 'search direction'. Thus, the new approximation is made by taking a 'step' in the search direction. That is, $\langle |\hat{f}_k^{(1)}|^2 \rangle$, the updated approximation to $\langle |\hat{h}_k|^2 \rangle$, could be given by

$$\langle |\hat{f}_k^{(1)}|^2 \rangle = \langle |\hat{f}_k^{(0)}|^2 \rangle + \Delta^{(0)} \hat{d}_k^{(0)}, \tag{8}$$

where $0 < \Delta^{(0)} \leq 1$ is the step size (independent of k), that is chosen small enough so that $\hat{d}_k^{(1)} = \langle |\hat{H}_k|^2 \rangle - \langle |\hat{F}_k^{(1)}|^2 \rangle$ is smaller (in some sense) than $\hat{d}_k^{(0)}$, and chosen large enough so that a reasonable rate of convergence is attained, where $\langle |\hat{F}_k^{(1)}|^2 \rangle$ is the energy spectrum of the field $F_n^{(1)}$ which is generated by exponentiating the Gaussian field $f_n^{(1)}$ with energy spectrum $\langle|\hat{f}_k^{(1)}|^2\rangle$. This idea is similar to the standard line search methods for finding extrema of nonlinear multivariate functions.

A series of such approximations leads to an iterative method:

$$\langle|\hat{f}_k^{(j+1)}|^2\rangle = \langle|\hat{f}_k^{(j)}|^2\rangle + \Delta^{(j)}\hat{d}_k^{(j)}, \tag{9}$$

where

$$\hat{d}_k^{(j)} = \langle|\hat{H}_k|^2\rangle - \langle|\hat{F}_k^{(j)}|^2\rangle. \tag{10}$$

For this method to succeed, it is not necessary for the linear approximation to be 'valid'; it is only necessary that the iterations converge. The method will work if the error $\hat{d}_k^{(j)}$ in the energy spectrum of the daughter log-normal field has the same sign as $\langle|\hat{h}_k|^2\rangle - \langle|\hat{f}_k^{(j)}|^2\rangle$, the error in the energy spectrum of the corresponding mother Gaussian field.

4. IMPLEMENTATION

Scaling log-normal simulations can be generated using the following algorithm. Some implementation issues are discussed below.

- Choose a first guess $\langle|\hat{f}_k^{(0)}|^2\rangle = c_f^2 k^{-\beta}$ of the ensemble average energy spectrum $\langle|h_n|^2\rangle$ for the final mother Gaussian field.

- Begin loop with $j = 0$:

1. Using the procedure described in Section 2, generate a Gaussian random field $f_n^{(j)}$ with the desired μ and σ^2, and with ensemble average energy spectrum $\langle|\hat{f}_k^{(j)}|^2\rangle$. In particular, set $E_k^{(t)} = C^2\langle|\hat{f}_k^{(j)}|^2\rangle$, where C^2 is a constant that is determined from the normalization condition (1).

2. Calculate $F_n^{(j)} = e^{f_n^{(j)}}$, and $|\hat{F}_k^{(j)}|^2$, the energy spectrum of $F_n^{(j)}$.

3. Approximate $\langle|\hat{F}_k^{(j)}|^2\rangle$, the ensemble average of $|\hat{F}_k^{(j)}|^2$ by performing a least-squares fit of a polynomial of degree l to the graph $\log_{10}\left(|\hat{F}_k^{(j)}|^2\right)$ vs. $\log_{10} k$. Note that if the energy spectrum were power-law, then this log-log plot would be linear (polynomial of degree 1). It was found that polynomials of degree 5 were both able to capture the deviations from power-law of $|\hat{F}_k^{(j)}|^2$, as well as, able to smooth out the fluctuations due to realization to realization variability.

4. Find the search direction $\hat{d}_k^{(j)} = \langle|\hat{H}_k|^2\rangle - \langle|\hat{F}_k^{(j)}|^2\rangle$, where $\langle|\hat{H}_k|^2\rangle = c_H^2 k^{-\beta}$ is the target energy spectrum (for the final log-normal simulation), β is the scale invariant exponent, and c_H is a constant (see below for the method used to determine c_H).

5. Set the energy spectrum for $f_n^{(j+1)}$, the next iterate Gaussian field: $\langle|\hat{f}_k^{(j+1)}|^2\rangle = \langle|\hat{f}_k^{(j)}|^2\rangle + \Delta^{(j)}\hat{d}_k^{(j)}$, where $0 < \Delta^{(j)} \leq 1$ is the step size.

6. If $\hat{d}_k^{(j)}$ is larger than some tolerance, set $j = j+1$, and go to step 1 of the loop. Otherwise, set $J = j$, and end loop.

- End loop.

- The final log-normal simulation is $F_n^{(J)}$ which has energy spectrum $\langle|\hat{F}_k^{(J)}|^2\rangle \approx c_H^2 k^{-\beta}$.

To avoid the technical issues involved in computing the theoretical value of the constant c_H, it is determined empirically. In particular, c_H is chosen from a linear fit to the graph $\log_{10}\left(|\hat{F}_k^{(j)}|^2\right)$ vs. $\log_{10} k$. This leads to increased stability of the method. Note that it is the normalization of the 'mother' Gaussian field, and not c_H, that is necessary for ensuring that the desired single-point statistics are obtained.

In the generation of the mother Gaussian fields, the same uncorrelated noise field ($g(x)$ of Section 2) should be used throughout. This ensures that 'realization to realization' fluctuations, that cause small variations in the estimated energy spectra, do not effect the convergence.

With a very simple determination of $\Delta^{(j)}$ and the first choice $\langle|\hat{f}_k^{(0)}|^2\rangle$ given above, it was found that the method converged for variances σ^2 up to 9. Due to issues associated with changes of magnitude, convergence was also slightly dependent on the value of μ. However, convergence was obtained for μ as high as 10. These ranges for the parameters were deemed to be sufficient for our purposes. However, if increased ranges are needed, more sophisticated means of choosing $\Delta^{(j)}$ and $\langle|\hat{f}_k^{(0)}|^2\rangle$ should lead to a significant increase in the the range of parameters for which convergence can be obtained.

As for the Gaussian fields generated using the procedure in Section 2, the method described in this section can be used to produce random fields of any dimension. Also, note that the random fields generated with Fourier methods are periodic. Thus, for more realistic simulations, one can subsample the original simulation (e.g. choose for the final simulation a part of size $N/4$ from the original field of size N).

Finally, although no tests have been performed, this method could be used to generate log-normal simulations that are not scaling. For instance, it may be possible to generate log-normal fields with multiple scaling regimes.

5. RESULTS

In this section, the results of a sample generation of a log-normal scaling simulation are presented. Figure 1 contains log-log plots of the energy spectra at the first iteration. It can be seen that the exponentiation of the Gaussian field has a roughening effect, i.e. the slope (on the log-log plot) of $|\hat{F}_k^{(0)}|^2$ is smaller than the slope β of

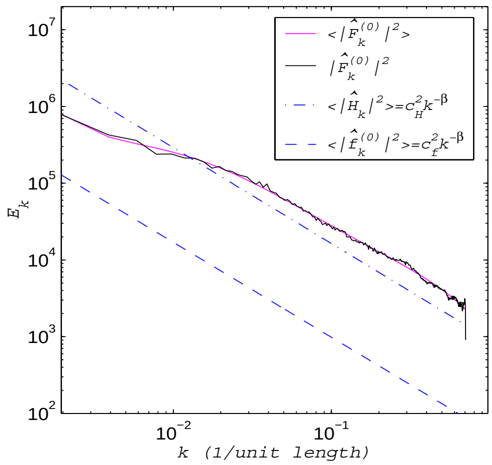

Figure 1: The energy spectra at the first iteration. The average energy spectrum $\langle|\hat{F}_k^{(0)}|^2\rangle$ is found by fitting a polynomial to the energy spectrum $|\hat{F}_k^{(0)}|^2$ of the first guess log-normal field. The straight lines are the target energy spectrum, $E_k^H = \langle|\hat{H}_k|^2\rangle$, and the first guess $\langle|\hat{f}_k^{(0)}|^2\rangle$.

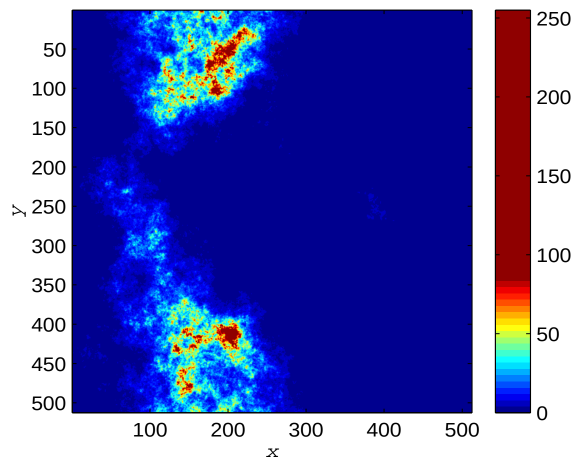

Figure 3: Two-dimensional discrete scaling log-normal simulation $F_n^{(J)}$ with mean $\mu = 0$, variance $\sigma^2 = 4$, and scale invariant exponent $\beta = 1.5$. The energy spectrum of this simulation is shown in Figure 2. A large variance was chosen to indicate that convergence could be obtained for variance values at least as large as could be expected for cloud data.

the target energy spectrum $E_k^H = \langle|\hat{H}_k|^2\rangle$. The most significant deviation of $|\hat{F}_k^{(0)}|^2$ from the target E_k^H is at small wave numbers; there is a more significant roughening at the large scales than at the small scales. Figure 2 shows the 'final' energy spectrum of the two-dimensional scaling log-normal simulation of size 512×512 that is shown in Figure 3. For additional tests of the generating technique, along with an implementation in Matlab, see Lewis and Austin (2002).

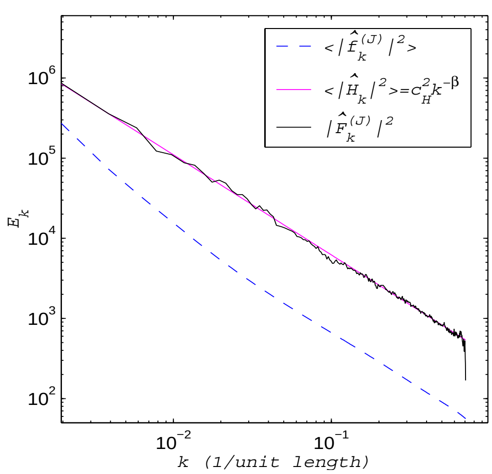

Figure 2: The energy spectra at convergence. The energy spectrum of the final simulation (see Figure 3) is $|\hat{F}_k^{(J)}|^2$. Also plotted are the target energy spectrum $\langle|\hat{H}_k|^2\rangle$ and the energy spectrum of the final Gaussian field $\langle|\hat{f}_k^{(J)}|^2\rangle$.

Acknowledgments The authors gratefully acknowledge support through funding of the Modeling of Clouds and Climate Proposal from the Natural Science and Engineering Research Council of Canada, the Canadian Foundation for Climate and Atmospheric Science, and the Meteorological Service of Canada.

References

Barker, H. W., B. A. Wielicki and L. Parker, 1996: A parameterization for computing grid-averaged solar fluxes for inhomogeneous marine boundary layer clouds. Part II: Validation using satellite data. *J. Atmos. Sci.*, **53**(16), 2304–2316.

Davis, A., A. Marshak, W. Wiscombe and R. Cahalan, 1994: Multifractal characterization of nonstationarity and intermittency in geophysical fields: Observed, retrieved, or simulated. *J. Geophys. Res.*, **99**, 8055–8072.

Evans, K. F., S. McFarlane and W. Wiscombe, 1999: Three-dimensional radiative transfer effects of two-point statistical representations of broken cloud structure. in *Proceedings of the 10th Conference on Atmospheric Radiation*, pp. 579–582. American Meteorological Society.

Frisch, U., 1995: *Turbulence*. Cambridge University Press, Cambridge, UK, 296pp.

Lewis, G. M. and P. H. Austin, 2002: An iterative method for generating scaling log-normal simulations. *http://www.eos.ubc.ca/research/clouds/radjp416*.

Lewis, G. M., P. H. Austin and M. Szczodrak, 2002: Spatial variability of marine boundary layer clouds. *submitted to J. Geophys. Res.*