JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 92, NO. D8, PAGES 9631–9643, AUGUST 20, 1987

A Space-Time Stochastic Model of Rainfall for Satellite Remote-Sensing Studies

THOMAS L. BELL

Laboratory for Atmospheres, NASA Goddard Space Flight Center, Greenbelt, Maryland

A model of the spatial and temporal distribution of rainfall is described that produces random spatial rainfall patterns with these characteristics: (1) The model is defined on a grid with each grid point representing the average rain rate over the surrounding grid box. (2) Rain occurs at any one grid point, on average, a specified percentage of the time and has a lognormal probability distribution. (3) Spatial correlation of the rainfall can be arbitrarily prescribed. (4) Time stepping is carried out so that large-scale features persist longer than small-scale features. Rain is generated in the model from the portion of a correlated Gaussian random field that exceeds a threshold. The portion of the field above the threshold is rescaled to have a lognormal probability distribution. Sample output of the model designed to mimic radar observations of rainfall during the Global Atmospheric Research Program Atlantic Tropical Experiment (GATE), is shown. The model is intended for use in evaluating sampling strategies for satellite remote-sensing of rainfall and for development of algorithms for converting radiant intensity received by an instrument from its field of view into rainfall amount.

1. INTRODUCTION

Satellite observation of rainfall is attractive because it promises information about rainfall rates on a nearly global basis, information that is unavailable from conventional earth-bound monitoring systems. It has, of course, its own set of problems. Rainfall rates must be indirectly inferred using remote-sensing methods, with all the errors associated with such techniques, and the observations are intermittent rather than continuous.

With these problems in mind, a satellite orbit must be selected that best satisfies the conflicting demands of being far from the earth, in order to see as much as possible, and of being close, in order to achieve high spatial resolution and to minimize the so-called "beam-filling problem." The beam-filling problem refers to the difficulty of relating the radiances received by a remote-sensing instrument from a spatially inhomogeneous but unresolved rain field within the field of view (FOV) of the instrument to the rainfall rate actually occurring within the FOV. It is discussed in more detail later in this section. As the instrument is brought closer to the surface and its FOV shrinks, the spatial variability of the rainfall within the FOV diminishes, and so does the beam-filling problem. At the same time, however, the satellite will view any one area of the globe less frequently and so will give a less-representative account of what the total rainfall in any one area has been, say, during a week or a month.

To evaluate the performance of the satellite and to optimize its parameters, it is helpful to have a model of how observed rainfall might vary in time and space. Such a model is described here, designed with the goal of addressing the two problems of beam-filling and the accuracy of area/time-averaged rainfall estimation from intermittent samples.

Let us first define the problem of estimating area/time-averaged rainfall, given the sampling characteristics of a satellite orbit. Suppose we wish to measure the average rainfall over an area A during a time period T. This may be written

$$R = \frac{1}{T}\int_0^T dt\,\frac{1}{A}\int_A d\mathbf{x}\; r(\mathbf{x}, t) \tag{1}$$

where $r(\mathbf{x}, t)$ is the rainfall rate at location $\mathbf{x}$ and time t, and where $\int_A$ denotes integration over the area A. (We arbitrarily call $t = 0$ the beginning of the period T.) Even with a perfect instrument capable of determining $r(\mathbf{x}, t)$ exactly, the satellite will pass over the area A only intermittently, at times t_i, $i = 1, \cdots, M$, during this period, and will sometimes observe only a portion A_i of the area A, so obtaining during one pass a subarea-averaged rainfall,

$$r_i = \frac{1}{A_i}\int_{A_i} d\mathbf{x}\; r(\mathbf{x}, t_i) \tag{2}$$

An estimate of R (equation (1)) using this instrument might be

$$\hat{R} = (1/M)\sum_{i=1}^{M} r_i \tag{3}$$

which is a special case of a more general weighted average of the observations,

$$\hat{R}_w = (1/M)\sum_{i=1}^{M} w_i r_i \tag{4}$$

where the weights w_i would be chosen to minimize the error

$$E^2 = \langle(\hat{R}_w - R)^2\rangle \tag{5}$$

The angle brackets denote an average over a hypothetical ensemble of rainfall events typical of the area A and time period T for which we are evaluating the satellite performance. The rainfall model could be used to suggest what the w_i in (4) should be. But even without this sophistication, one would want to know the typical size of the error E for the estimate (3). This error is the best one could expect of the satellite, neglecting the additional inaccuracies in remote-sensing methods of measuring rain rates.

Suppose we know the covariance matrix of the rainfall,

$$U(\mathbf{x}, t; \mathbf{x}', t') = \langle r(\mathbf{x}, t) r(\mathbf{x}', t')\rangle - \langle r(\mathbf{x}, t)\rangle\langle r(\mathbf{x}', t')\rangle \tag{6}$$

Paper number 7D0321.

We can then write an explicit expression for the error (5) which is a function of the covariance U alone, since the error, when expressions (1) and (3) are substituted into it, is quadratic in $r(\mathbf{x}, t)$, and U contains all information about the second-order moment statistics of r. (We must also assume that the statistics of the rainfall are reasonably homogeneous in space and time. A diurnal cycle in the rainfall statistics, for example, might introduce additional errors in our estimate (3)). Thus for a rainfall model to estimate well the errors one should expect in $\hat{R}$ obtained from a satellite, the model statistics must at the very least reproduce the covariance U in (6). The model described in this paper can, in principal, capture this aspect of rainfall exactly, assuming of course that one can provide a plausible form for the covariance structure U in the first place. This can only come from analyses of rainfall data from different climatic regimes. The example given here is designed to mimic aspects of rainfall data observed in GATE. (GATE is an acronym for the Global Atmosphere Research Program Atlantic Tropical Experiment.)

The beam-filling problem requires information of a different sort. For this problem we need to know how variable rainfall rate might be within a single FOV of the instrument. This is needed in order to explore the relationship between the average rainfall occurring in a FOV,

$$R_{\mathrm{FOV}} = A_{\mathrm{FOV}}^{-1} \int_{A_{\mathrm{FOV}}} d\mathbf{x}\, r(\mathbf{x}, t) \tag{7}$$

and the radiance emanating from the FOV received by the instrument,

$$I_{\mathrm{FOV}} = A_{\mathrm{FOV}}^{-1} \int_{A_{\mathrm{FOV}}} d\mathbf{x}\, \mathscr{I}[r(\mathbf{x}, t)] \tag{8}$$

where $\mathscr{I}$ represents the radiant intensity (at some frequency) produced by the rainfall at point $\mathbf{x}$. (In reality, $\mathscr{I}$ will be determined not only by the local rainfall rate, but also by such factors as vertical column structure of the raining clouds and instrumental noise, and so has a "random" component.) The problem of relating R_{FOV} to I_{FOV} when $\mathscr{I}$ is nonlinear in $r(\mathbf{x}, t)$ is referred to as the beam-filling problem. R_{FOV} is what we want to know, but I_{FOV} is what we measure. For fixed R_{FOV}, I_{FOV} can vary depending on the spatial variability in $r(\mathbf{x}, t)$. A model is needed to generate typical distributions of $r(\mathbf{x}, t)$ within a FOV to investigate the beam-filling problem and to develop algorithms to convert I_{FOV} to R_{FOV} with as little error as possible.

The model is described in the sections that follow. The closest analogous two-dimensional rainfall model the author is aware of is that of *Mejia and Rodriguez-Iturbe* [1974*b*], which was further developed by *Bras and Rodriguez-Iturbe* [1976]. The version they describe generates rainfall with a nearly normal probability distribution, although it could be modified using a technique similar to what is used here to generate rainfall with other probability distributions. The model described in this paper also takes advantage of a technique (the "fast Fourier transform" method) that contributes significantly to its numerical efficiency. This is important in carrying out some of the studies described above, since simulations of the order of a year's worth of data with a time step of a fraction of an hour are envisioned. Other stochastic models which should be mentioned are those explored by *Waymire et al.* [1984], which attempt to capture the development of rainbands and rain cells and are related in form to a model proposed by *Le Cam* [1961]; and those of *Lovejoy and Mandelbrot* [1985], *Lovejoy and Schertzer* [1985], and *Schertzer and Lovejoy* [this issue], whose models exhibit "scaling" or "fractal" behavior. These latter models are discussed further in section 2.

Section 2 describes the constraints the rainfall model was designed to satisfy. Section 3 gives a simplified account of how the model is constructed to satisfy those constraints. Section 4 describes some of the characteristics and output of a numerical model. Rain fields on a 256×256 grid, each grid point representing rain rate in a 4×4 km^2 area, are generated. Some possible modifications of the model appropriate to various sampling studies are described in section 5. A detailed, technical discussion of the model is given in the appendix.

2. CONSTRAINTS ON THE MODEL

The rainfall model is constructed with several constraints in mind, foremost among them that the space-time covariance structure of rainfall be as accurately represented as possible, since, as was argued in section 1, the covariance statistics of rainfall largely govern the sampling errors inherent in an observational system. Some of the constraints serve to simplify the model so that only a few parameters are needed to specify completely its behavior and can be established using rainfall data. Some of the constraints can be relaxed, at the cost of requiring more parameters to be specified.

First let us list the constraints, and then discuss the reasons why they are emphasized. The constraints are as follows:

1. Rain fields on a finite (bounded) area will be simulated. Each grid point of the model represents rain falling in a grid square centered on that point.
2. The statistics of the rainfield are independent of spatial position and direction in the field (i.e., they are spatially homogeneous and isotropic).
3. It rains, on average, only a fraction f of the time at any one grid point. When it rains, the rain rate r has probability density $p(r)$. This probability will be taken to be lognormal here.
4. The spatial correlation between any two points $\mathbf{x}$ and $\mathbf{y}$ in the field will be specified and, because of assumption 2, will depend only on $|\mathbf{x}-\mathbf{y}|$.
5. Area-averaged rainfall will have time-lagged autocorrelations that depend on the size of the area. Larger areas will have longer correlation times.

When the model is used to investigate the effects of intermittent sampling of the rainfall within a square area of perhaps 500 km on a side at some location on the earth, assumption 2 of spatial homogeneity of the statistics is a plausible first approximation, at least over large oceanic regions where little conventional rainfall data is available and where satellites may be the only sources of information. The assumption may be less justified in areas where strong topographic effects or persistent thermal contrasts exist. The assumption of statistical isotropy for the rain fields must also be viewed as an approximation. *Zawadzki* [1973], for instance, gives examples of anisotropy of spatial correlations for a storm over Montreal, Canada. Neglect of anisotropy will mean that the model will not capture certain subtle differences in how the statistics of satellite observations of rainfall may depend on the orientation of the satellite's nadir track with respect to the equator or on the shape of the FOV. An extension of the model to include anisotropic statistics could easily be developed, but at the costs of additional complexity and more parameters.

Assumption 3 raises the issue of choosing the probability distribution that best represents instantaneous rain rates.

Many different statistical characterizations of rainfall have been explored, and lognormality is a frequent candidate, but a large fraction of these characterizations have involved quantities such as the volume of rain falling during a storm (see *Biondini* [1976], for example), or the daily rain amount (see *Crow et al.* [1979], for example), or temporal or areal extent (see *Houze and Cheng* [1977] and *Lopez* [1977], for example). Rain gages give time-averaged rain rates over an area of the order of 10^{-8} km^2, much smaller than the areas whose statistics must be represented by the model, though spatial coherency of rainfall and time averaging may justify extrapolating the data to larger areas. Such data have, of course, been extensively analyzed. See *Drufuca and Zawadzki* [1975], for just one instance. Near-instantaneous (1-min average) rain rates from rain gages have been carefully studied by *Jones and Sims* [1978].

The model discussed here requires the statistics of instantaneous rainfall averaged over areas of the order of 1 km^2, typical of the resolution of radar data. The lognormal distribution is sometimes recommended as a convenient representation for the probability distribution of radar-derived rain rate (see *Crane* [1985, 1986], for instance). It is easy to convince oneself that the probability of large rain rates does not decrease with rain rate nearly as fast as a Gaussian probability distribution would predict. The probability of large values of a variable that is distributed lognormally decreases (asymptotically) more slowly than an exponential, yet faster than any power of the variable; however, over a finite but considerable range of the variable, its probability distribution can be difficult to distinguish from an algebraic (or "hyperbolic") falloff (r^{-p}) [see *Montroll and Schlesinger*, 1982], especially when one is trying to estimate the distribution from a limited amount of data. *Lovejoy and Mandelbrot* [1985] have proposed that high rain rates may in fact be hyperbolically distributed, as one manifestation of "scaling" behavior in rain. It is unfortunately not easy to construct a clean statistical test of which of the two distributions describes rain better, because enormous amounts of data are needed to assess the behavior of the tail of the probability distribution on which the hypotheses focus, and the data correlations in space and time are difficult to take into account with most tests. Some careful statistical work is needed in this area.

Here rain rate has been generated with a lognormal probability distribution, since the distribution describes rain rates reasonably well, with two parameters, over the range of values observed in the GATE radar data sets whose statistics we are trying to imitate (B. Kedem, L. S. Chiu, and G. R. North, Estimation of mean rain rate: Application to satellite observations, submitted to the *Journal of Geophysical Research*, 1987; hereinafter referred to as KCN). Furthermore, the distribution permits us to discuss spatial covariances such as (6) in a conventional way, whereas moments of distributions with a hyperbolic tail require considerably more care. The alternative proposed by *Lovejoy and Mandelbrot* [1985] is, however, extremely interesting. *Lovejoy and Mandelbrot* [1985], *Lovejoy and Schertzer* [1985], and *Schertzer and Lovejoy* [this issue] have described models which may be capable of reproducing spatial and temporal rainfall statistics (but with hyperbolic-tail distributions). Further study is needed to see what the implications of their models may be for the remote sensing of rainfall.

Finally, let us turn briefly to constraints 4 and 5. These constraints assume that spatial and temporal correlations of rainfall over the scales of interest can be obtained from radar data. An example of this will be explored in detail in section 4.

3. Model Description

To generate time-dependent rain fields with the characteristics described in section 2, extensive use is made of the possibility of generating random fields $g(\mathbf{x})$ that have Gaussian statistics at each grid point and are spatially correlated, by the well-known device of representing the field as a Fourier series,

$$g(\mathbf{x}) = \sum_{\mathbf{k}} a_{\mathbf{k}} \exp(i\mathbf{k}^T\mathbf{x}) \tag{9}$$

where superscript T denotes matrix transpose. See the appendix for a more precise description of the summation range (equations (A2) and (A3), in particular). The random coefficients $a_{\mathbf{k}}$ have real and imaginary parts $(a_{\mathbf{k}})_R$, $(a_{\mathbf{k}})_I$ that are normally distributed and uncorrelated with each other and with all the other coefficients and whose variances are chosen to produce the desired spatial correlation. Details of this technique are discussed in the appendix. Let us assume, then, that we are able to generate a random Gaussian field with spatial correlation

$$\langle g(\mathbf{x})g(\mathbf{y})\rangle = c_g(|\mathbf{x}-\mathbf{y}|) \tag{10}$$

We also assume that g at any point has mean 0 and variance 1. The field $g(\mathbf{x})$ might be related physically to the vertical wind field or to an index of vertical stability of the lower atmosphere, although these should be viewed more as helpful analogies than the beginnings of a physical theory.

We convert this field to a rainfield $r(\mathbf{x})$ by finding a transformation $\mathscr{R}$,

$$r = \mathscr{R}(g) \tag{11}$$

with the property that the probability distribution of r is the one given in constraint 3 in the previous section; that is, r should be 0 a fraction $1-f$ of the time and distributed as $p(r)$ the rest of the time. Figure 1 illustrates the way that this is accomplished. A threshold g_0 is found such that the Gaussian field g exceeds the threshold a fraction f of the time. The portion of the field that exceeds the threshold is then rescaled so that its probability distribution is $p(r)$. This can be accomplished as follows:

The variable g is first converted to a variable u that is uniformly distributed between 0 and 1, using the transformation

$$u = G(g) \tag{12}$$

$$G(g) = (2\pi)^{-1/2}\int_{-\infty}^{g} \exp(-\tfrac{1}{2}y^2)\,dy \tag{13}$$

so that the probability density ρ of u is

$$\rho(u) = 1 \qquad 0 < u < 1 \tag{14}$$

To insure that it rains only a fraction f of the time, set

$$r = 0 \qquad u < 1-f \tag{15}$$

Thus the threshold g_0 is the solution of the equation

$$G(g_0) = 1-f \tag{16}$$

Values of u greater than $1-f$ are converted to rainfall, using the complement of the cumulative distribution function for $p(r)$,

$$\mathscr{C}(r) = \int_r^{\infty} p(s)\,ds \tag{17}$$

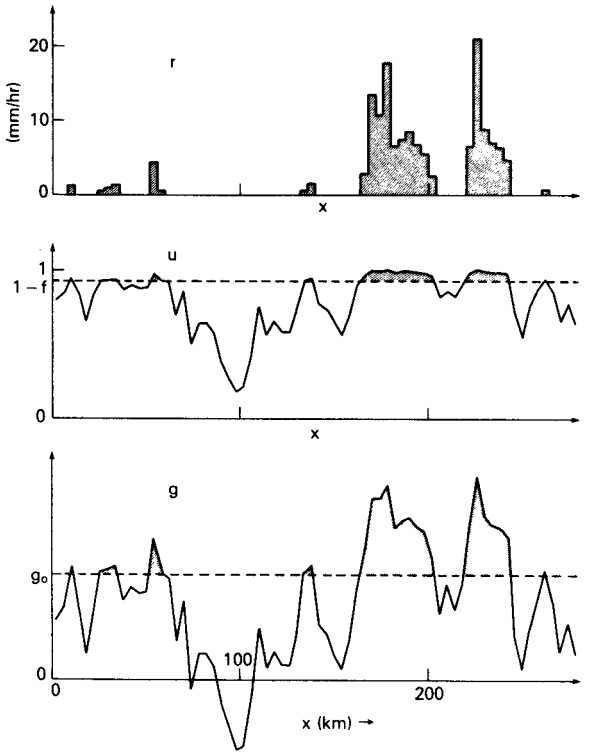

Fig. 1. Transformation of Gaussian field $g(\mathbf{x})$ to rain field $r(\mathbf{x})$. The field g exceeds the threshold g_0 a fraction f of the time. It is transformed to variable u with a uniform probability distribution. Variable u is converted to lognormally distributed r whenever it exceeds $1 - f$.

by writing

$$r = \mathscr{C}^{-1}[(1 - u)/f] \qquad u > 1 - f \tag{18}$$

where $\mathscr{C}^{-1}$ means the functional inverse of $\mathscr{C}$. The probability distribution of r (conditional on $r > 0$) defined this way will then be $p(r)$. The transformation $\mathscr{R}$ in (11) can thus be written

$$\begin{aligned} \mathscr{R}(g) &= 0 & g \le g_0 \\ \mathscr{R}(g) &= \mathscr{C}^{-1}\{[1 - G(g)]/f\} & g > g_0 \end{aligned} \tag{19}$$

Constraints 2 and 3 are now satisfied. To satisfy constraint 4, we need to obtain a rain field with a spatial correlation structure

$$\frac{\langle r'(\mathbf{x})r'(\mathbf{y})\rangle}{\text{var } (r)} = c_r(|\mathbf{x} - \mathbf{y}|) \tag{20}$$

Here primes denote deviations from the mean $\langle r\rangle$ and var (r) denotes the variance of r, $\langle (r')^2\rangle$. To do this we need to take into account the effect the transformation $\mathscr{R}$ has on the Gaussian field correlation $c_g(|\mathbf{x} - \mathbf{y}|)$ defined in (10). Suppose two Gaussian variables g and h, both with mean 0 and variance 1, have correlation c:

$$c = \text{corr } (g, h) \tag{21}$$

$$\text{corr } (g, h) \equiv \langle (g - \langle g\rangle)(h - \langle h\rangle)\rangle/[\text{var } (g) \text{ var } (h)]^{1/2} \tag{22}$$

The corresponding rain rates generated by the transformation $\mathscr{R}$ have correlation

$$c_r = \text{corr } [\mathscr{R}(g), \mathscr{R}(h)] \tag{23}$$

Equation (23) defines a relation between the correlation c of the Gaussian variables and the corresponding correlation of these variables converted to rain rates, which we can write symbolically as

$$c_r = H(c) \tag{24}$$

An expansion of $H(c)$ in powers of c can be obtained using Mehler's formula (see *Slepian* [1972], for example), in which the coefficients in the expansion involve squares of the expectation value of products of $\mathscr{R}(g)$ and the Hermite polynomials $H_k(g)$. (I thank M. S. Taqqu for drawing this device to my attention.)

Analytical expressions for H can be obtained for certain special cases of the transformation $\mathscr{R}$ in (11). In particular, when the Gaussian variable g is converted to a lognormally distributed variable using $r = \exp(\mu + \sigma g)$, *Matalas* [1967] and *Mejia and Rodriguez-Iturbe* [1974a] obtain

$$H(c) = [\exp(c\sigma^2) - 1]/[\exp(\sigma^2) - 1] \tag{25}$$

Although the functional form of H may not always be easy to obtain analytically, it can accurately be estimated numerically by a procedure whose details are given in the appendix. Given H, we are able to satisfy constraint 4, by choosing the correlation of the Gaussian field to have the form

$$c_g(|\mathbf{x} - \mathbf{y}|) = H^{-1}[c_r(|\mathbf{x} - \mathbf{y}|)] \tag{26}$$

where $c_r(|\mathbf{x} - \mathbf{y}|)$ is assumed given in constraint 4.

Troublesome possibilities exist that the relation (24) might not be one-to-one, in which case the inverse H^{-1} would be ambiguous, or that no value of c might generate a desired value of c_r (as occurs in example (25) for c_r near -1), in which case H^{-1} would be undefined. These theoretical possibilities have not yet been encountered in attempting to model rainfall data, but the modeler should be aware of them. The first seems unlikely if $\mathscr{R}$(equation (11)) is monotonic. Large negative correlations could lead to the second problem, but they do not seem to occur with real data: a lower bound for negative correlations of the nonnegative variable r is $-\langle r\rangle^2/\text{var } (r)$, and rain rate variance seems generally to be much larger than the squared mean rainfall $\langle r\rangle^2$.

We turn finally to constraint 5. Constraint 5 is satisfied by giving the coefficients $a_{\mathbf{k}}$ in (9) time dependence, with larger spatial scales (small $|\mathbf{k}|$) evolving more slowly than small scales (large $|\mathbf{k}|$), but arranged such that the rainfall probability distribution continues to satisfy constraints 2–4. Let the real and imaginary parts of the coefficients $a_{\mathbf{k}}(t)$ evolve in time, so that their lagged correlations decrease exponentially with lag τ,

$$\text{corr } \{[a_{\mathbf{k}}(t + \tau)]_{R,I}, [a_{\mathbf{k}}(t)]_{R,I}\} = \exp(-|\tau|/\tau_{\mathbf{k}}) \tag{27}$$

This is accomplished by letting the coefficients $a_{\mathbf{k}}(t)$ satisfy Markov equations in time. A review of the properties of such equations may be found in the work by *Jenkins and Watts* [1968]. Evolution of $a_{\mathbf{k}}(t)$ in time occurs with time steps of size Δt,

$$a_{\mathbf{k}}(t) = \beta_{\mathbf{k}} a_{\mathbf{k}}(t - \Delta t) + z_{\mathbf{k}}(t) \tag{28}$$

The $z_{\mathbf{k}}(t)$ are random Gaussian variables whose real and imaginary components are uncorrelated with each other or with themselves at different times (white noise), or with $z_{\mathbf{k}'}(t)$, $\mathbf{k}' \neq \mathbf{k}$.

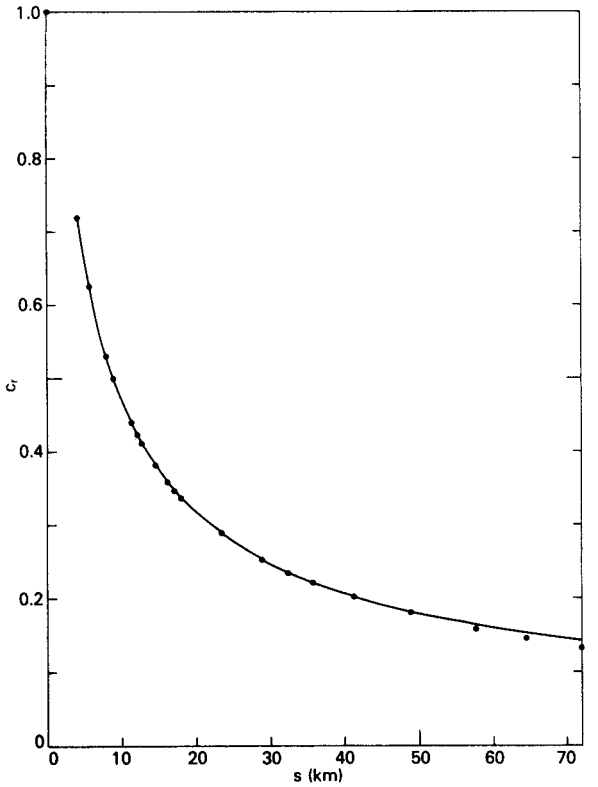

Fig. 2. The spatial correlation of GATE I rainfall (courtesy of A. McConnell, private communication, 1985). Sampling errors in correlations are estimated to be of the order of ± 0.1 for large separation s.

The real numbers $\beta_{\mathbf{k}}$ represent the lag-Δt autocorrelation of $a_{\mathbf{k}}(t)$ (see equation 27),

$$\beta_{\mathbf{k}} = \exp\left[-\Delta t/\tau_{\mathbf{k}}\right] \tag{29}$$

If the variance of $z_{\mathbf{k}}$ is set at

$$\text{var}\left[(z_{\mathbf{k}})_{R,I}\right] = (1 - \beta_{\mathbf{k}}^2)\ \text{var}\left[(a_{\mathbf{k}})_{R,I}\right] \tag{30}$$

then they generate $a_{\mathbf{k}}$ with just the variance needed to assure that the field $g(\mathbf{x})$ in (9) has the desired spatial correlations (equation (10)).

The time constants $\tau_{\mathbf{k}}$ in (29) must still be specified. To satisfy constraint 5, they must decrease with increasing $|\mathbf{k}|$. Determining their precise values from the empirically estimated time correlations of rainfall, however, is not so easy. In principle, one could use the lagged correlations of rainfall at a grid point to obtain the corresponding lagged correlations required of the Gaussian field g at a grid point, using (24), and then choose $\tau_{\mathbf{k}}$ to generate the right lagged correlations in g. A less methodical route to choosing the $\tau_{\mathbf{k}}$ has been taken, which will be described in section 4.

Because a Markov process has been used (equation (28)) to model time development, lagged correlations of $r(\mathbf{x}, t)$ will decay approximately exponentially with time ("approximately" because rainfall at a grid point is determined by a sum of many Fourier components with different characteristic time scales and because transformation (11) will distort the correlations). Other statistical models of the time development are possible, from higher-order autoregressive processes [see *Jenkins and Watts*, 1968] to "fractional Brownian motion" processes, such as are described by *Mandelbrot* [1971] and *Chi et al.* [1973], as long as the constraint is met that the process reproduce the required variance of the coefficients $a_{\mathbf{k}}$ (cf. equation (30)).

Overall advection of the rainfield can be introduced quite easily in this model by using the Fourier representation of an advecting field (cf. equation (9))

$$g(\mathbf{x} - \mathbf{U}t) = \sum_{\mathbf{k}} a_{\mathbf{k}} \exp\left(-i\mathbf{k}^T\mathbf{U}t\right) \exp\left(i\mathbf{k}^T\mathbf{x}\right) \tag{31}$$

where $\mathbf{U}$ is the advection velocity of the field. By changing the $\beta_{\mathbf{k}}$ in (28) to complex numbers,

$$\beta_{\mathbf{k}} = \exp\left[-(\Delta t/\tau_{\mathbf{k}}) - i\mathbf{k}^T\mathbf{U}\Delta t\right] \tag{32}$$

generation of new fields at each time step with (28) can include advection of the field as well. Taylor's hypothesis, as applied to rainfields [*Zawadski*, 1973], would presumably work accurately in this model for times small compared to the smallest correlation time $\tau_{\mathbf{k}}$ of spatial scales with a significant amount of variability. See *Gupta and Waymire*, [this issue] for a recent discussion of Taylor's hypothesis for other stochastic models.

4. Sample Model Output

A numerical version of the model was constructed that incorporates statistical features of quarter-hourly radar observations of rainfall during the GATE I period (June 28 to July 16, 1974). These statistical features were obtained by analyzing data in a rainfall atlas covering a 400-km-diameter area with a 4×4 km^2 grid, centered on 8°30'N latitude and 23°30'W longitude off the west coast of Africa. The methods used in producing this unique atlas are described by *Hudlow and Patterson* [1979].

The probability distribution for rainfall during this period has been analyzed by KCN. They find that approximately 8% of the 4×4 km^2 boxes have rain in them and that the probability distribution of rain rate is nearly lognormal. Accordingly, the following values have been assigned for the rain rate probability distribution parameters:

$$f = 0.08 \tag{33a}$$

$$\mu = \langle \ln r \rangle = 1.14 \tag{33b}$$

$$\sigma^2 = \text{var}(\ln r) = 1.21 \tag{33c}$$

which are typical of the values obtained by KCN. The parameters μ and σ are the mean and standard deviation of the normal probability distribution for $\ln r$, when r is expressed in units of millimeters per hour.

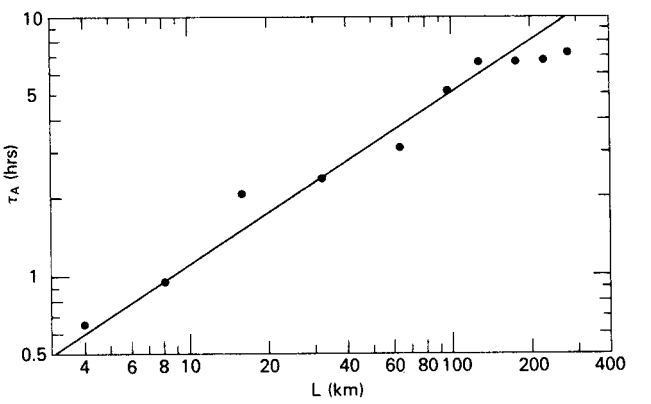

Fig. 3. The correlation times of GATE I rainfall, averaged over $L \times L$ area, from a study by *Laughlin* [1981].

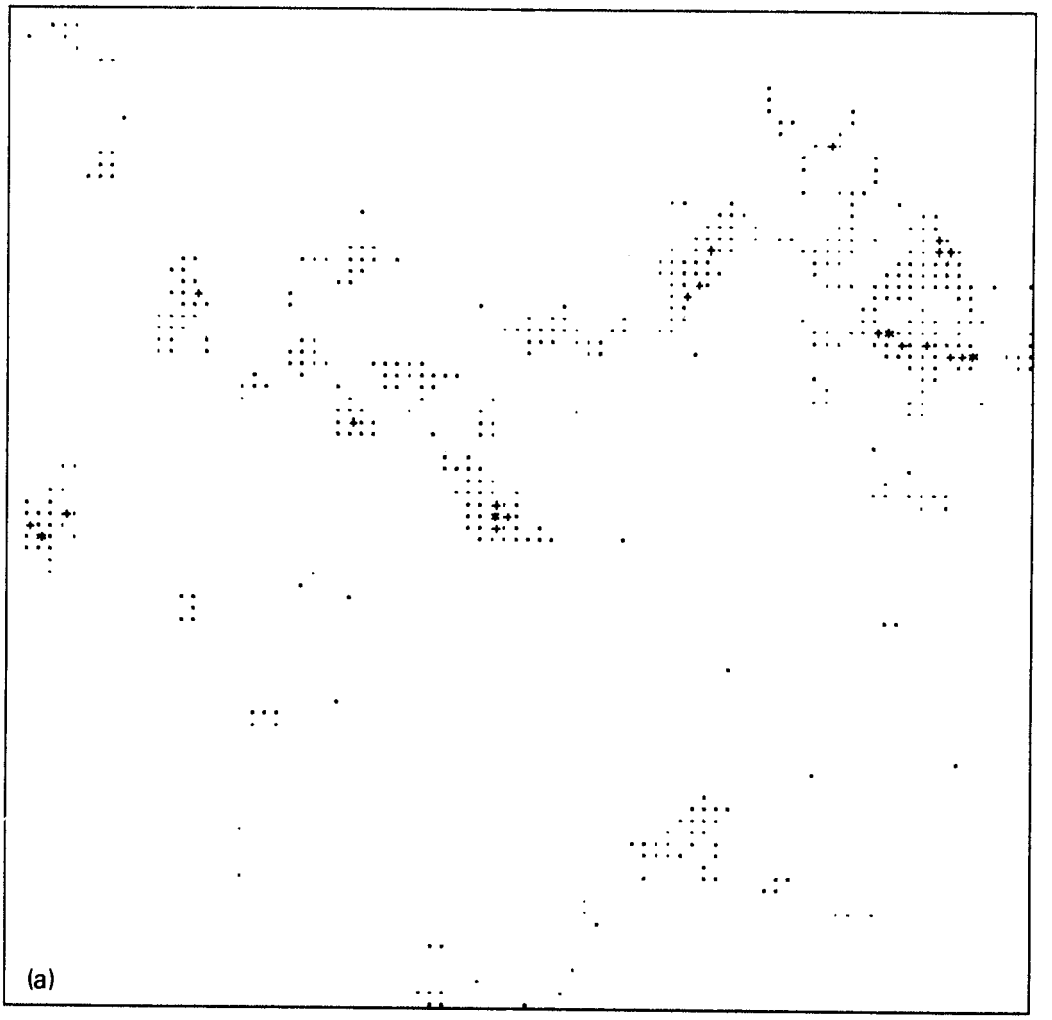

Fig. 4. Model rainfield, 85 × 85 grid point portion from a 256 × 256 grid point simulation. Dots indicate rainfield rate $r < R_1 = 10.5$ mm/hr, pluses indicate $R_1 < r < R_2 = 22.0$ mm/hr, and asterisks indicate $r > R_2$. The divisions R_1 and R_2 are chosen such that half of the total rain (lognormally distributed, with parameters as in (equation (33)) is due to rain rates less than R_1, and one quarter is due to rain rates greater than R_2. (*a*) Rain occurs over 6.6% of area. (*b*) Rainy area is equal to 19.9%.

The spatial correlation of rain rate was derived (A. McConnell, private communication, 1985) from the covariance of the rainfall

$$U(|\mathbf{x} - \mathbf{y}|) = \langle [r(\mathbf{x}) - \langle r(\mathbf{x})\rangle][r(\mathbf{y}) - \langle r(\mathbf{y})\rangle]\rangle \qquad (34)$$

where the angle brackets denote an average over all GATE I observations and over all pairs of points separated by distance $|\mathbf{x} - \mathbf{y}|$. The spatial correlation defined in (20) is then estimated from

$$c_r(|\mathbf{x} - \mathbf{y}|) = U(|\mathbf{x} - \mathbf{y}|)/U(\mathbf{0}) \qquad (35)$$

A plot of this quantity versus separation is shown in Figure 2, out to a separation of 72 km. The data is described well by the formula (A. McConnell, private communication, 1985)

$$c_r(s) = 1/(\tfrac{1}{4}s + 0.63682)^{2/3} \qquad s \geq 4 \text{ km} \qquad (36)$$

(displayed as a smooth curve in Figure 2), and this empirical fit has been used to specify the spatial correlation of rain generated by the model. Note that although the correlation (36) is written in terms of the continuous variable s, it is, in fact, only valid for separations on a 4-km grid. Moreover, it should not be interpreted as the correlation of "point" rainfall; rather, it represents the correlation of rainfall averaged over 4-km squares.

How accurate are the empirical correlations on which (36) is based? In particular, how much might the estimate (36) of the correlation vary from the "true" correlations one would obtain from a much larger data set? It is not easy to determine the sampling error in these correlations because the data from which they are derived are so dependent. A rough order of magnitude estimate of the error can, however, be obtained from the following considerations: The covariance of rainfall separated by a distance $\mathbf{s}$ is the average of the products of rain rate at all pairs of points separated by $\mathbf{s}$, as expressed in (34). For separations $|\mathbf{s}|$ around 72 km, the accuracy of the estimate is measured by its standard deviation,

$$\sigma_U = N_e^{-1/2}\{\text{var } [r(\mathbf{x})] \text{ var } [r(\mathbf{y})]\}^{1/2}$$

where N_e is the effective number of independent samples entering into the average, and the fact is used that at these larger separations the rainfall at the two points are roughly independent. But what is N_e? Because rain is correlated in space, the number of spatially independent samples is far fewer than the number of grid points averaged over. The variance of area-

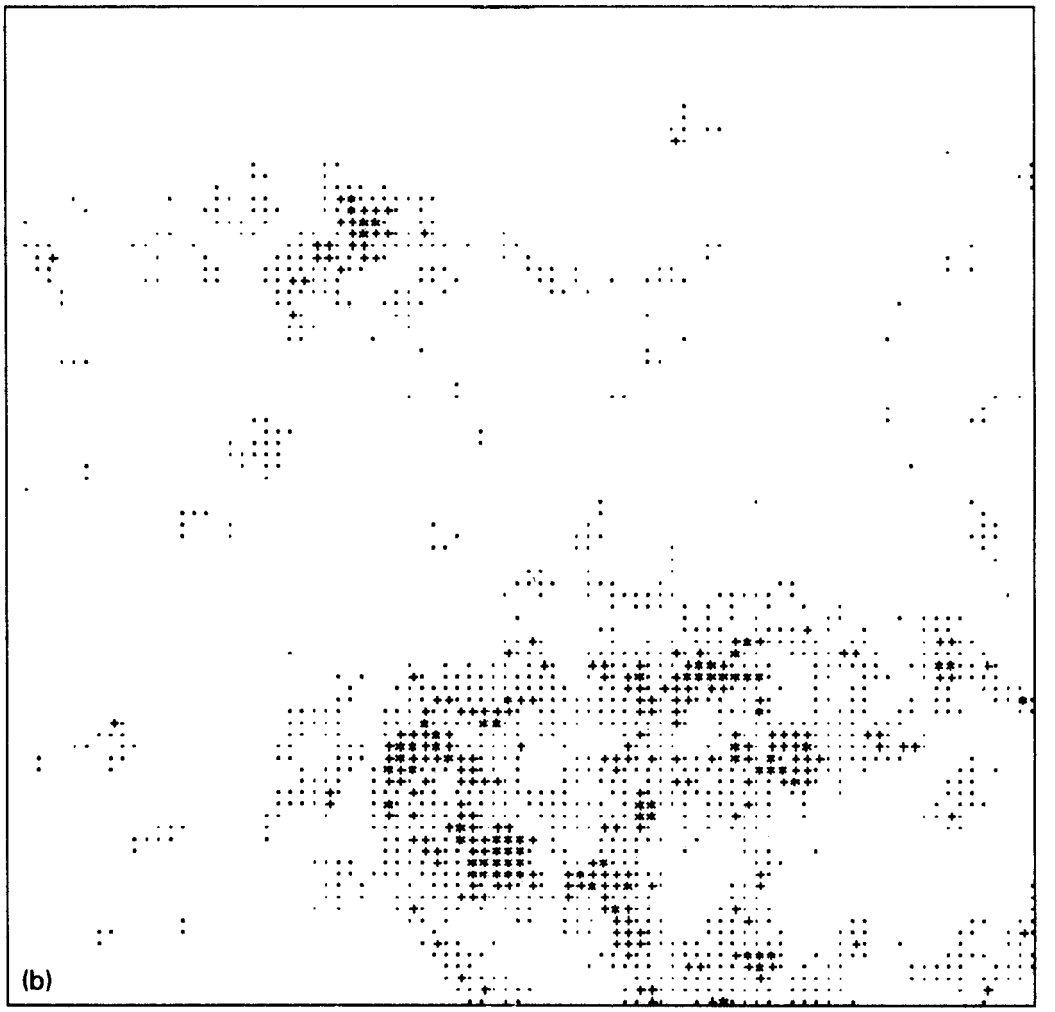

Fig. 4. (continued)

averaged rainfall in GATE I decreases very nearly as $A^{-1/3}$ with area A, as has been found by *Laughlin* [1981], instead of as A^{-1}, which would be the case if there were a finite correlation length. The effective number of independent samples in spatial averages is therefore estimated over the data as

$$[A/(4 \times 4 \text{ km}^2)]^{1/3}$$

where $A/(4 \times 4 \text{ km}^2)$ is the number of pairs of grid points used in the average (approximately 5200, or about two thirds of the 400-km-diameter GATE area, for averages over points separated by $s = 72$ km with a fixed orientation). The data is also correlated in time. As discussed later in this section, the correlation time of area-averaged rainfall for the largest area in GATE is of the order of $\tau = 7$ hours. Using *Leith*'s [1973] estimate of the independent sample time, 2τ, one would conclude that there are $18 \times 24/2\tau$ temporally independent samples in the 18 days of GATE data used in the average. Multiplying the number of independent spatial and temporal samples together, an estimate

$$N_e \approx (5200)^{1/3} \times 31$$

$$N_e = 540$$

is obtained which is far, far fewer than a naive estimate based on the product of the number of grid points in the GATE area times the number of snapshots would suggest ($\sim 10^7$). Correlations in the neighborhood of $s = 72$ km are thus estimated to be accurate to within two standard deviations, $2\sigma_c \approx 2N_e^{-1/2} = 0.09$. (The small additional contribution to the sampling error from estimating var(r), which appears in the denominator of the expression (35) for the correlation, is ignored). This sampling error may at best be regarded as an order of magnitude estimate of the true error. It might be slightly smaller because correlations for two orientations of the vector $\mathbf{s}$ were averaged together to obtain the values in Figure 2. On the other hand, during the GATE I period there is a marked tendency for disturbances to concentrate in the southeastern half of the GATE area [see *Hudlow and Patterson*, 1979], so that the assumption of spatial homogeneity of the statistics can only be approximately valid, and the assumption of isotropy, as already discussed, is also only approximate. Analysis of the 16 days of radar data available from the GATE II period (July 28 to August 15, 1974), for which disturbances were weaker but more evenly distributed over the GATE area, gives spatial correlations consistently less than those in Figure 2, by slightly less than 0.1 (A. McConnell, private communication, 1985). It therefore seems reasonable to trust the spatial correlations represented by (36) to perhaps ± 0.1.

Another possible concern is that of extrapolating (36) out to

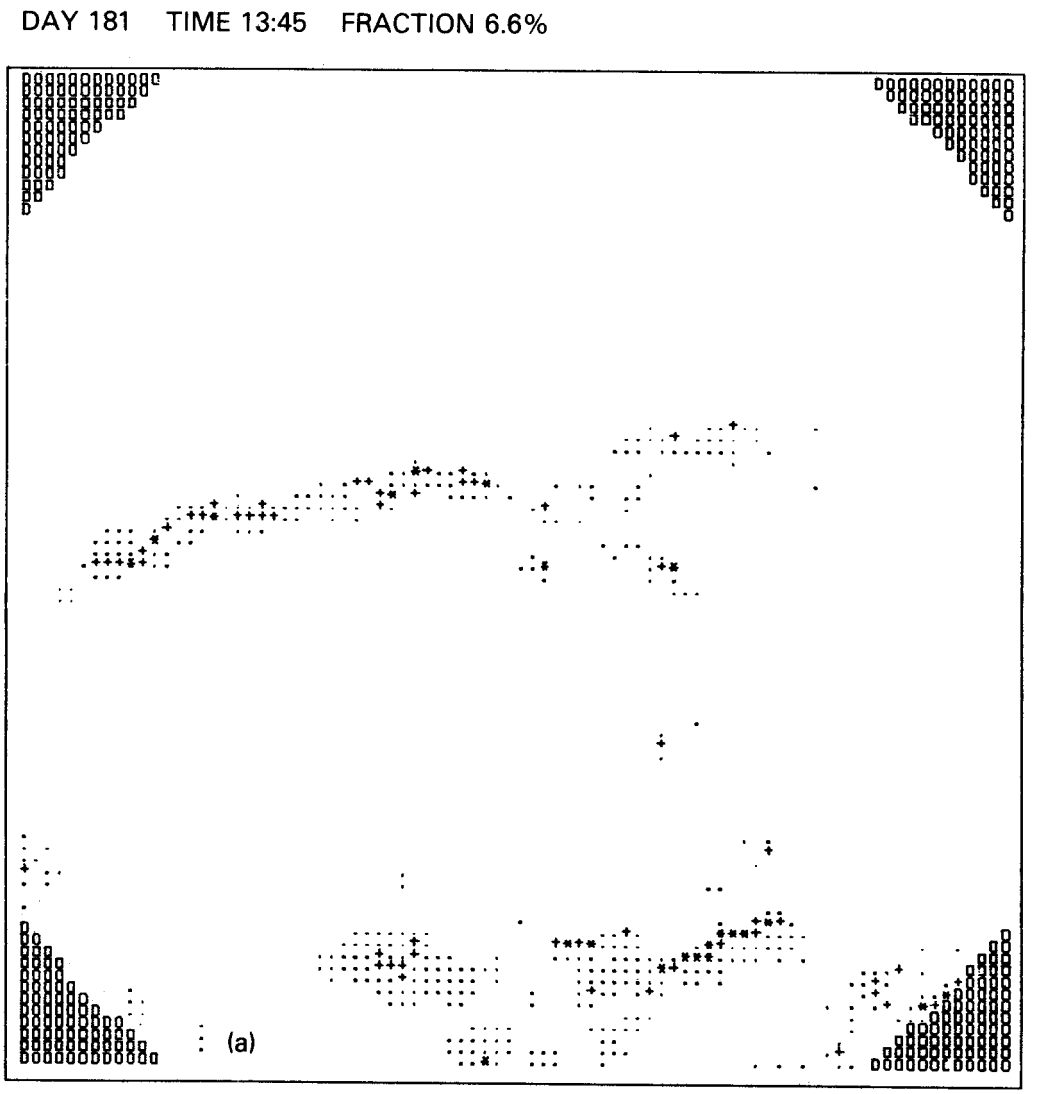

Fig. 5. Some sample GATE rainfall images, using the same plotting convention as in Figure 4. No data is available in the corners of the image. (*a*) Rain occurs over 6.6% of area (not including the corners). (*b*) Rainy area is equal to 19.9%.

the distances employed in these simulations (~ 500 km). The fact that the variance of area-averaged rainfall decreases very nearly as $A^{-1/3}$ up to areas as large as 280×280 km^2 (found for GATE I data by *Laughlin* [1981]), which would be predicted by the power law behavior of (36), reinforces using (36) as a reasonable extrapolation out to these distances. A faster decrease at large separation s cannot be ruled out, however, because of the ± 0.1 uncertainty in the empirical correlations.

The temporal behavior of the model is dictated by the coefficients $\beta_{\mathbf{k}}$ in (28), which are, in turn, fixed by the correlation times $\tau_{\mathbf{k}}$ defined in (29). To determine these correlation times, results reported by *Laughlin* [1981] were used to suggest appropriate time scales for the various wave numbers. *Laughlin* [1981] obtained lagged correlations for rainfall rate averaged over areas A ranging from 4×4 km^2 to 280×280 km^2 and fit them to a form $\exp(-\tau/\tau_A)$ where τ is lag and τ_A is the correlation time. His results have been graphed for τ_A versus L ($A = L^2$) in Figure 3, and fitted to a form

$$\tau_A(L) = c_\tau L^{2/3} \tag{37}$$

with $c_\tau = 0.24$ when τ_A is expressed in units of hours and L is stated in kilometers. If we associate fluctuations of rainfall averaged over an $L \times L$ area with fluctuations of Fourier components with wavelength $2L$, then we can associate time scales with Fourier components by equating

$$\tau_{\mathbf{k}} = \tau_A(L = \pi/k) = c_\tau(\pi/k)^{2/3} \tag{38}$$

The form of (38) is suggested by scaling arguments for turbulent fluids frequently invoked to explain the Kolmogorov kinetic energy spectrum for turbulent fluids (see *Orszag* [1977] for a review of these arguments) and by the observed spectrum of winds and temperatures in the troposphere apparently following the Kolmogorov spectrum up to the scales of interest in this model. (A review of these results may be found in the work by *Lilly* [1983]). The time scale for $\mathbf{k} = 0$ is not defined by (38) and is set equal to the time scale of the smallest nonzero wave number.

For simulations of areas of the order of 512×512 km^2, (38) would imply a time constant of the order of 15 hours, using the constant c_τ from the fit to Figure 3. Because the time scales in the data for the largest scales seem somewhat less than the formula (37) would give, the largest time scale in the model has been somewhat arbitrarily reduced to 12 hours, and a power law form (38) has been assumed for all other scales, so

Fig. 5. (continued)

that the smallest time scale of the model, corresponding to spatial scales of the order of 4 km, is 0.5 hours. However, it should be noted that the time scales discussed here are correlation times for the Gaussian field $g(\mathbf{x})$ from which the rainfield is generated by (11), rather than for the rainfield itself, whose correlations tend to be smaller than for the Gaussian field (see Figure 9 in the appendix). Further investigation of the amount of temporal correlation in the Gaussian field needed to reproduce the rainfield temporal statistics is desirable.

Some examples of the kinds of rainfall intensity patterns generated by the model on a 256 × 256 grid, with each grid point representing rainfield with a 4 × 4 km^2 area, are shown in Figures 4*a* and 4*b*. Only a portion of the area is shown, 85 × 85 in size. For comparison, some scenes from GATE are shown in Figures 5*a* and 5*b* chosen to have similar fractional area covered by rain. The most striking difference in the model images and the radar images is that the spatial organization of the radar image is stronger. The model has more small-scale activity and does not generate elongated structures like rainbands so frequently. There may also be a tendency for heavier rain to be associated with larger-scale rain events than may be the case in the GATE data.

The model was tested to see if it produces a rainfall probability distribution similar to GATE and to see if its spatial covariance is similar to GATE, and the agreement was excellent, as it must be (if the numerical implementation is correct), since the model is forced to have those characteristics. The correlation time for rain in a grid box (i.e., averaged over a 4-km square) was found to be approximately 0.5 hour, for rain averaged over a 64-km square it was 3 hours, for a 512-km square it was 8 hours. Although this last correlation time is less than (37) predicts, it has been (somewhat arbitrarily) decided not to force it to a larger value, since the maximum correlation time supported by the empirical data graphed in Figure 3 is 8 hours.

An interesting feature of rainfall patterns is that the "diameters" of precipitating areas tend to be lognormally distributed [*Lopez*, 1976; 1977]. Because the model rainfall patterns have more diffuse boundaries than do radar images, it is not as straightforward to assign precipitating areas a precise diameter as it is with the radar images. As an alternative method of analyzing the distribution of rainy areas, a one-dimensional version of such an analysis has been carried out. Straight lines have been drawn through the images. Rainy segments of various lengths occur along such a line. The distribution of

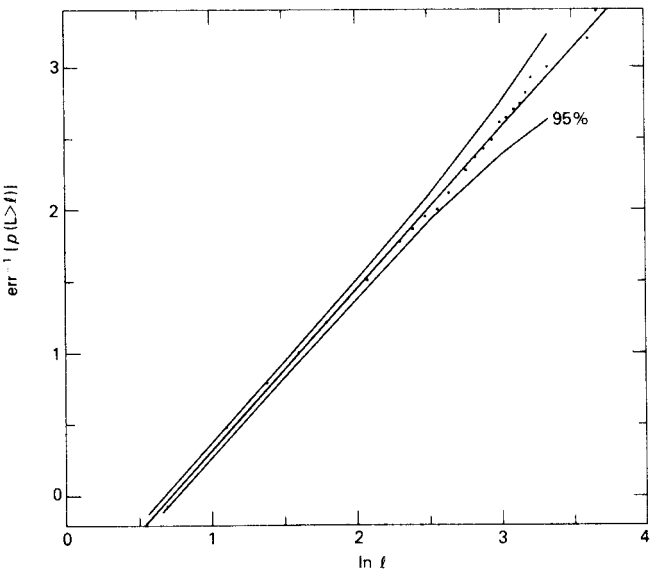

Fig. 6. Test for lognormality of the distribution of lengths of rainy segments along an arbitrary line drawn through model rainfields. Points show the inverse Gaussian error function $\mathrm{err}^{-1}(p)$ where p is the cumulative probability distribution for the lengths, $p(L > l)$. A total of 2907 segments were counted in the analysis. A straight line would indicate perfect lognormality. Points should lie within the 95% confidence limits. The mean of ln l is 0.715 and its standard deviation is 0.893, where l is measured in units of grid points.

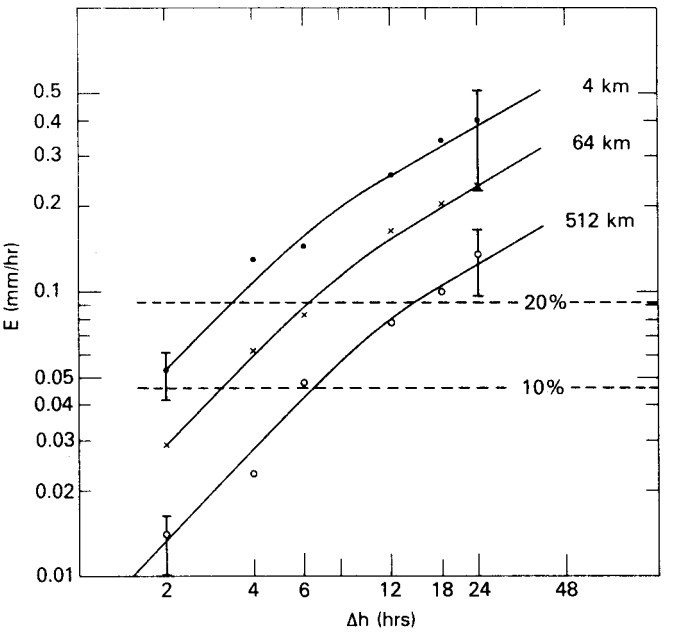

Fig. 7. Root-mean-square errors in estimates of total monthly rainfall falling on lines of length 4, 64, or 512 km, as a function of the time between observations of the rainfall. The dashed lines indicate the size of the error as a percentage of the climatic mean rainfall for the simulation, $\langle r \rangle = 0.46$ mm/hr. Representative error bars in the values of E are shown, based on the limited number (60) of sample months used to estimate E.

rainy segment lengths has been obtained by analyzing many rainfall model images in this way. The distribution of these lengths is very close to lognormal, as shown in Figure 6 by plotting the inverse error function of the cumulative distribution of rainy segment lengths l versus ln l. The points lie very close to a straight line and are nearly all within the 95% confidence limits for a lognormal distribution.

As an example of the use to which this kind of model can be put, a simplified version of one of the questions the model is intended to address is considered. Suppose that a satellite observes an area at equally spaced intervals Δh for a month. How accurate an estimate for the total rainfall during the month will the satellite obtain, assuming that the entire area is observed each time and that the observations themselves are perfectly accurate (i.e., neglecting instrumental and remote-sensing errors)? In other words, what is the value of E in (6)?

A one-dimensional version of the model has been used which produces rainfall patterns exactly like those that occur along a line in the two-dimensional model. Figure 7 shows values of E for line-averaged rainfall of length 4, 64, and 512 km, as a function of sampling time interval Δh. These values are based on averages over 60 months of hourly data generated by the model. Error bars (95% confidence) indicate the estimated accuracy of the values of E. For reference, dashed lines show the ratio of E to the mean rainfall rate of the model for the parameters chosen in (9), 0.46 mm/hr. The error increases with the sampling interval Δh, of course, and decreases with the size of the area averaged over.

5. Discussion

The model described here can reproduce the spatial covariance and grid point statistics of a rain field exactly and can approximate the temporal statistics. For some purposes this is all that is required of a statistical model. Construction of the model requires knowing the spatial and temporal statistics of the rain fields being modeled. In this we are fortunate in having the GATE radar data set. More analyses of radar and rain gage data will be required to extend the model to other climatic regimes. A higher spatial resolution model is desirable for beam-filling studies, which requires extrapolating the statistics of the GATE data to this resolution, or analyzing the statistics of higher resolution radar data.

The model as described here is isotropic and does not generate rainfall patterns as spatially well defined as radar images of rainbands. An anisotropic covariance structure could be used that would generate elliptically shaped patterns oriented along a chosen direction, but this would only partially mimic the phenomenon.

Other possible modifications of the model might include modulation of the parameters with time, to represent a diurnal or seasonal cycle; use of higher-order autoregressive or even fractional Brownian processes, to produce smoother time evolution of the field; and imposition of spatial inhomogeneities in the statistics, such as a tendency to rain more in part of the simulation field. All of this, of course, would be accomplished at the cost of additional complexity in the model and the need for more parameters to be specified.

Appendix: Details of Model Construction

A more detailed discussion of the model and a description of some of the numerical methods used follows.

Generating $g(\mathbf{x})$

Central to the model is the generation of a correlated Gaussian field $g(\mathbf{x})$. The set of grid points for which the field is specified will be a $N \times N$ square array, where N is even; in fact, N should ideally be a power of 2 ($N = 2^K$) in order to

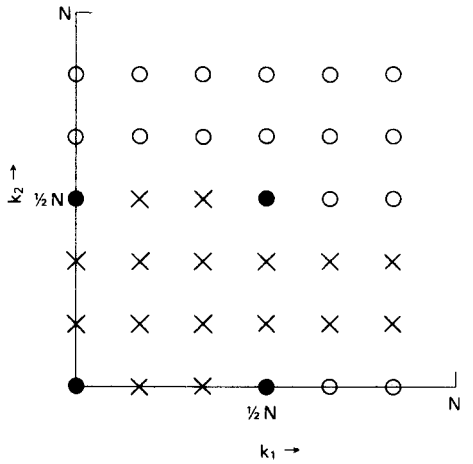

Fig. 8. The constraints (A15) places on the coefficients a_k for $N = 6$. The four solid points must be real, the points marked with crosses are complex, and the points marked with circles are related by (A15) to ones marked by a cross and are thus not independent.

take maximum advantage of fast Fourier transform algorithms generally available on computers as FORTRAN-callable subroutines. The locations of the grid points $\mathbf{x}$ are thus given by

$$\mathbf{x} = (x_1, x_2) \qquad 0 \le x_1, x_2 \le N - 1 \tag{A1}$$

where x_1 and x_2 are integers.

The Fourier expansion of $g(\mathbf{x})$ is written

$$g(\mathbf{x}) = \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} a_{\mathbf{k}} u_{\mathbf{k}}(\mathbf{x}) \tag{A2}$$

$$u_{\mathbf{k}}(\mathbf{x}) = \exp(i2\pi \mathbf{k}^T \mathbf{x}/N) \tag{A3}$$

(where, again, T denotes matrix transpose). Equation (A2) can be inverted to give the coefficients

$$a_{\mathbf{k}} = N^{-2} \sum_{\mathbf{x}} g(\mathbf{x}) u_{\mathbf{k}}^*(\mathbf{x}) \tag{A4}$$

$$\sum_{\mathbf{x}} \equiv \sum_{x_1=0}^{N-1} \sum_{x_2=0}^{N-1} \tag{A5}$$

where asterisks indicate complex conjugation. By using the expansion (A2) we are implicitly assuming that $g(\mathbf{x})$ is periodic, because

$$u_{\mathbf{k}}(\mathbf{x}) = u_{\mathbf{k}}(\mathbf{x} + \mathbf{d}) \tag{A6}$$

for all vectors

$$\mathbf{d} = (n_1 N, n_2 N) \tag{A7}$$

n_1, n_2 arbitrary integers. Consequently,

$$g(\mathbf{x}) = g(\mathbf{x} + \mathbf{d}) \tag{A8}$$

and opposite edges of the field behave as if they are joined. By using (10) and (A4), one can show that the covariance of the coefficients satisfies the equations

$$\langle a_{\mathbf{k}} a_{\mathbf{l}}^* \rangle = 0 \qquad \mathbf{k} \ne \mathbf{l} \tag{A9}$$

$$\langle a_{\mathbf{k}}^* a_{\mathbf{k}} \rangle = N^{-2} \sum_{\mathbf{x}} c_g(|\mathbf{x}|) u_{\mathbf{k}}^*(\mathbf{x}) \tag{A10}$$

which can be inverted to give the correlation function in terms of the variances,

$$c_g(|\mathbf{x}|) = \sum_{\mathbf{k}} \langle a_{\mathbf{k}}^* a_{\mathbf{k}} \rangle u_{\mathbf{k}}(\mathbf{x}) \tag{A11}$$

Equations (A6) and (A11) require that c_g reflect the periodicity of the field in satisfying

$$\begin{aligned} c_g(|\mathbf{x}|) &\equiv c_g(x_1, x_2) \\ c_g(x_1, x_2) &= c_g(N - x_1, x_2) = c_g(x_1, N - x_2) \end{aligned} \tag{A12}$$

Thus if the model is to be used to simulate a rain field in which spatial correlations decrease with increasing distance, the periodic field generated by this procedure must be divided into four quadrants, any one of which will contain a field whose correlations decrease as desired right out to the edges of the quadrant. If a larger area of the model is used, it will have the peculiarity that as the (apparent) separation between two points becomes greater than $N/2$ in either direction, their correlation starts increasing again, according to (A12).

The coefficients $a_{\mathbf{k}}$ are complex, and at first sight it appears that there are $2N^2$ numbers being used in (A2) to generate N^2 values of the field $g(\mathbf{x})$. However, the $a_{\mathbf{k}}$ are constrained by the requirement that $g(\mathbf{x})$ is real,

$$g(\mathbf{x}) = g^*(\mathbf{x}) \tag{A13}$$

and this, in addition to the periodicity of $u_{\mathbf{k}}(\mathbf{x})$ in $\mathbf{k}$,

$$u_{\mathbf{k}+\mathbf{d}}(\mathbf{x}) = u_{\mathbf{k}}(\mathbf{x}) \tag{A14}$$

where $\mathbf{d}$ is given by (A7), implies that the $a_{\mathbf{k}}$ satisfy

$$a_{\mathbf{k}} = a_{\mathbf{d}-\mathbf{k}}^* \tag{A15}$$

The constraints this relation implies for $a_{\mathbf{k}}$ are illustrated in Figure 8, for $N = 6$. The four points marked with a solid circle are constrained by (A15) to be real. They occur at

$$\mathbf{k} = (0, 0), (0, N/2), (N/2, 0), (N/2, N/2) \tag{A16}$$

Each coefficient marked with an open circle is constrained by (A15) to be the complex conjugate of one of the coefficients marked by crosses in Figure 8. The total number of independent coefficients is N^2, as it should be.

To generate a random field $g(\mathbf{x})$ then, we choose independent random Gaussian-distributed numbers and assign them to each coefficient $a_{\mathbf{k}}$ (real and imaginary parts, where necessary) as follows: Define

$$\sigma_{\mathbf{k}}^2 = h_{\mathbf{k}} \langle a_{\mathbf{k}}^* a_{\mathbf{k}} \rangle \tag{A17}$$

where $\langle a_{\mathbf{k}}^* a_{\mathbf{k}} \rangle$ is given by (A10) and $h_{\mathbf{k}} = 1/2$ except for the four vectors $\mathbf{k}$ in (A16), for which $a_{\mathbf{k}}$ is real (marked by solid dots in Figure 8) and for which $h_{\mathbf{k}} = 1$. The coefficients $a_{\mathbf{k}}$ are then given the values

$$a_{\mathbf{k}} = \sigma_{\mathbf{k}} e_{\mathbf{k}} \tag{A18}$$

where $e_{\mathbf{k}}$ is a random complex number (except for $\mathbf{k}$ such that $a_{\mathbf{k}}$ is real, in which cases the imaginary part of $e_{\mathbf{k}}$ is zero); the $e_{\mathbf{k}}$ have mean 0 and variance 1 (real and imaginary parts) and are uncorrelated with one another but satisfy (A15). In practice, this means assigning all of the crossed points random values and filling in the values of the circled points using constraint (A15).

The field $g(\mathbf{x})$ generated from (A2) will, when averaged over many samples, have the desired spatial covariance $c_g(|\mathbf{x} - \mathbf{y}|)$, with the constraint described in (A12) due to periodicity of the domain. Although it might appear that Gaussian fields with arbitrarily specified covariance structure $c_g(|\mathbf{x}|)$ may be generated by (A2), this is not in fact the case, because (A10) requires

that the Fourier transform of the spatial correlation be nonnegative. Since the form of $c_r(|\mathbf{x}|)$ is derived from actual rainfall data, it will have a nonnegative Fourier transform (although careless handling of missing data in a time series can sometimes cause trouble). However, it is not obvious that the corresponding Gaussian field covariance $c_g(|\mathbf{x}|) = H^{-1}[c_r(|\mathbf{x}|)]$ (equation (26)) might not have a negative Fourier transform. Negative Fourier transforms can also be generated when the periodicity constraints (equations (A12)) are imposed on the correlation structure. The most natural way to satisfy (A12) is to use the empirically derived correlations $c_r(|\mathbf{x}|)$ (transformed by equation (26)) out to the boundaries of a quadrant of the simulation field (i.e., for $x_1, x_2 \leq N/2$)), and for apparent separations $\mathbf{x}$ beyond the quadrant to assign the spatial correlation the values dictated by the relations in (A12). However, this produces a "kink" in the correlation structure at the quadrant boundaries, which can generate negative Fourier components of the resulting $c_g(|\mathbf{x}|)$. This problem has been met with in simulating very small rain fields ($N \leq 16$), where a few large-$|\mathbf{k}|$ Fourier components proved to be negative but very small. (It was not met with in the larger simulations.) One solution to these problems is simply to replace the negative Fourier components with $\sigma_{\mathbf{k}}^2 = 0$ or with the absolute value of the Fourier component. The model will then generate rainfields with spatial correlations different from the empirical values originally attempted. The size of this difference can be determined by inverting the modified spectrum using (A11). However, this difference may be within the statistical uncertainty of the empirical estimates and so may be quite acceptable. As already mentioned, these problems were not met with in the large-area simulations whose results are presented here.

Obtaining $\mathscr{R}$ for Lognormally Distributed Rainfall

The conversion of g to lognormal rain rates in (11) can be made numerically simpler than what is implied by (19). Since the function

$$\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty \exp(-t^2)\, dt \tag{A19}$$

is available as a FORTRAN-callable function on many computer systems, it is convenient to use

$$u = 1 - G(g) \tag{A20}$$

instead of (12) to generate a uniformly distributed variable, since this becomes

$$u = \tfrac{1}{2} \operatorname{erfc}(g/\sqrt{2}) \tag{A21}$$

in terms of erfc. Values of u in the region

$$f \leq u \leq 1 \tag{A22}$$

are converted to 0 rain rate, and the portion $0 < u < f$ is used to generate rain. To generate lognormal rain rates, the device of writing

$$r = \exp(\mu + \sigma\xi) \tag{A23}$$

is used, where μ is the desired mean of $\ln(r)$ and σ^2 is the desired variance of $\ln(r)$, and ξ is Gaussian-distributed, with mean 0 and variance 1. Values of ξ are generated from u by writing

$$\xi = G^{-1}(1 - u/f) \qquad 0 < u < f \tag{A24}$$

which can be written as

$$\xi = \sqrt{2}\, \operatorname{erf}^{-1}(1 - 2u/f) \qquad 0 < u < f \tag{A25}$$

where

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-t^2)\, dt \tag{A26}$$

FORTRAN-callable subroutines for erf^{-1} are available in the IMSL subroutine library. We thus have a relatively simple numerical procedure to convert $g(\mathbf{x})$ to a rain field, with rain occurring on average a fraction f of the time, and with a lognormal distribution characterized by the parameters μ and σ.

Obtaining $c_g(|\mathbf{x}|)$

The last ingredient of the model to be discussed is how to obtain $c_g(|\mathbf{x}|)$, since it is $c_r(|\mathbf{x}|)$ (equation (20)) that is specified. This requires computing (23) (written symbolically as equation (24)) in as economical a way as possible, and inverting it to obtain c_g in terms of c_r. This can be done numerically, and a description of a procedure developed to handle this problem follows:

Given two correlated Gaussian variables g and h, a multivariate probability distribution for them can be written

$$P(g, h) = \frac{1}{2\pi} \frac{1}{\sqrt{1 - c^2}} \exp\left[-\tfrac{1}{2}(g^2 + h^2 - 2cgh)/(1 - c^2)\right] \tag{A27}$$

Variables g and h have means 0 and variances 1 and correlation c. They are converted to rain rates using $\mathscr{R}$,

$$r_g = \mathscr{R}(g) \qquad r_h = \mathscr{R}(h) \tag{A28}$$

and the average of the product of these rain rates is given by

$$\begin{aligned} \langle r_g r_h \rangle &= \int_{-\infty}^{\infty} dg \int_{-\infty}^{\infty} dh\, P(g, h)\, r_g r_h \\ &= \int_{g_0}^{\infty} dg \int_{g_0}^{\infty} dh\, P(g, h)\, r_g r_h \end{aligned} \tag{A29}$$

where (19) has been used in order to include only those values of g, $h > g_0$ corresponding to nonzero rain rate. Since the transformation $\mathscr{R}$ in (A28) is determined once f, μ, and σ^2 are fixed (see previous section), (A29) can be evaluated numerically for a range of values of c in (A27). The quantity $\langle r_g r_h \rangle$ can then be converted to the associated c_r defined in (23), given the mean and variance of r, using the lognormal results

$$\langle r \rangle = f \exp(\mu + \tfrac{1}{2}\sigma^2) \tag{A30}$$

$$\langle r^2 \rangle = f \exp(2\mu + 2\sigma^2) \tag{A31}$$

For c near 1 the probability distribution $P(g, h)$ in (A27) becomes highly peaked, and it is convenient for numerical integration to use as integration variables u and v, with

$$g = 2^{-1/2}(u + v) \tag{A32a}$$

$$h = 2^{-1/2}(u - v) \tag{A32b}$$

so that

$$\begin{aligned} \langle r_g r_h \rangle = 2 \int_{\sqrt{2}g_0}^{\infty} du \int_0^{u - \sqrt{2}g_0} & p_u(u) p_v(v) \\ & \cdot \mathscr{R}[2^{-1/2}(u + v)] \mathscr{R}[2^{-1/2}(u - v)] \end{aligned} \tag{A33}$$

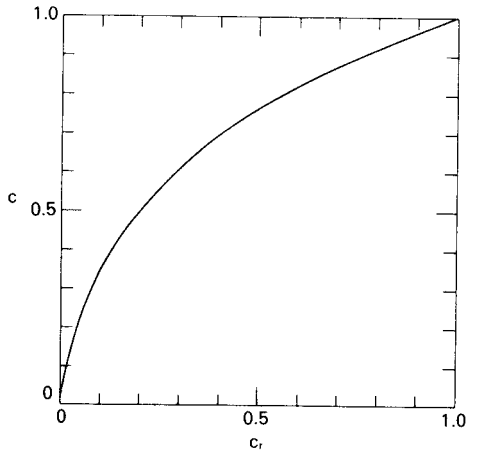

Fig. 9. An example of the relationship of c to c_r expressed by (23) for a particular choice of the rainfall statistics typical of the GATE rainfall, for which $f = 0.08$, $\mu = 1.14$, and $\sigma = 1.1$. Values of $\langle r_g r_h \rangle$ were obtained following the numerical method outlined in the appendix (see section on obtaining $c_g(|\mathbf{x}|)$) for values of $c = 0.2$, 0.4, 0.6, and 0.8, and the curve plotted here was obtained by cubic-spline interpolation.

with

$$p_u(u) = [2\pi(1 + c)]^{-1/2} \exp [-\tfrac{1}{2}u^2/(1 + c)] \quad \text{(A34)}$$

$$p_v(v) = [2\pi(1 - c)]^{-1/2} \exp [-\tfrac{1}{2}v^2/(1 - c)] \quad \text{(A35)}$$

Equations (A28) and (A32) are used to make the dependence of r_g and r_h on u and v explicit in (A33). For c near 1, $p_v(v)$ is much narrower than $p_u(u)$, and so numerical integration of variable v should use narrower steps than those for variable u.

An example of the integration results for parameter values $f = 0.08$, $\mu = 1.14$, and $\sigma = 1.1$, characteristic of GATE radar rainfall data, is shown in Figure 9. Once the relationship between c and c_r is known for a representative set of values of c, it is a simple matter to convert any given value of c_r to the corresponding value of c_g, using standard interpolation methods. This solves the last problem in creating a numerically efficient realization of the model sketched in section 3.

Acknowledgments. I have benefited greatly from numerous conversations with A. Abdullah, L. S. Chiu, A. McConnell, R. Martin, G. R. North, and D. A. Short. I wish particularly to thank G. R. North for his encouragement, A. McConnell and D. S. Short for permission to use unpublished results from their analyses of the GATE data, and A. Abdullah for assistance with some of the numerical work. Suggestions by M. S. Taqqu for improving the text are gratefully acknowledged.

REFERENCES

Biondini, R., Cloud motion and rainfall statistics, *J. Appl. Meteorol.*, *15*, 205–224, 1976.

Bras, R. L., and I. Rodríguez-Iturbe, Rainfall generation: A nonstationary time-varying multidimensional model, *Water Resour. Res.*, *12*, 450–456, 1976.

Chi, M., E. Neal, and G. K. Young, Practical application of fractional Brownian motion and noise to synthetic hydrology, *Water Resour. Res.*, *9*, 1523–1533, 1973.

Crane, R. K., Evaluation of global and CCIR models for estimation of rain rate statistics, *Radio Sci.*, *20*, 865–879, 1985.

Crane, R. K., Horizontal small scale structure of precipitation, in *Proceedings of the 23rd Conference on Radar Meteorology, Snowmass, Colorado*, vol. 1, pp. 181–184, American Meteorological Society, Boston, Mass., 1986.

Crow, E. L., A. B. Long, J. E. Dye, and A. J. Heymsfield, Results of a randomized hail suppression experiment in Northeast Colorado, II, Surface data base and primary statistical analysis, *J. Appl. Meteorol.*, *18*, 1538–1558, 1979.

Drufuca, G., and I. I. Zawadzki, Statistics of raingage data, *J. Appl. Meteorol.*, *14*, 1419–1429, 1975.

Gupta, V. K., and E. Waymire, On Taylor's hypothesis and dissipation in rainfall, *J. Geophys. Res.*, this issue.

Houze, R. A. and C.-P. Cheng, Radar characteristics of tropical convection observed during GATE: Mean properties and trends over the summer season, *Mon. Weather Rev.*, *105*, 964–980, 1977.

Hudlow, M. D., and V. L. Patterson, GATE radar rainfall atlas, *NOAA Spec. Rep.*, 158 pp., Natl. Oceanic and Atmos. Admin., Boulder, Colo., 1979.

Jenkins, G. M., and D. G. Watts, *Spectral Analysis and Its Applications*, chap. 5, Holden-Day, Oakland, Calif., 1967.

Jones, D. M. A., and A. L. Sims, Climatology of instantaneous rainfall rates, *J. Appl. Meteorol.*, *17*, 1135–1140, 1978.

Laughlin, C. R., On the effect of temporal sampling on the observation of mean rainfall, Precipitation Measurements From Space, Workshop report, edited by D. Atlas and O. W. Thiele, *NASA publ.*, Goddard Space Flight Center, Greenbelt, Md., 1981.

Le Cam, L., A stochastic description of precipitation, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. III, edited by J. Neyman, pp. 165–186, University of California Press, Berkeley, 1961.

Leith, C. E., The standard error of time-average estimates of climatic means, *J. Appl. Meteorol.*, *12*, 1066–1069, 1973.

Lilly, D. K., Mesoscale variability of the atmosphere, in *Mesoscale Meteorology—Theories, Observations and Models*, edited by D. K. Lilly and T. Gal-Chen, pp. 13–24, D. Reidel, Hingham, Mass., 1983.

Lopez, R. E., Radar characteristics of the cloud populations of tropical disturbances in the northwest Atlantic, *Mon. Weather Rev.*, *104*, 269–283, 1976.

Lopez, R. E., The lognormal distribution and cumulus cloud populations, *Mon. Weather Rev.*, *105*, 865–872, 1977.

Lovejoy, S., and B. B. Mandelbrot, Fractal properties of rain, and a fractal model, *Tellus*, *37A*, 209–232, 1985.

Lovejoy, S., and D. Schertzer, Generalized scale invariance in the atmosphere and fractal models of rain, *Water Resour. Res.*, *21*, 1233–1250, 1985.

Mandelbrot, B. B., A fast fractional Gaussian noise generator, *Water Resour. Res.*, *7*, 543–553, 1971.

Matalas, N. C., Mathematical assessment of synthetic hydrology, *Water Resour. Res.*, *3*, 937–945, 1967.

Mejia, J. M., and I. Rodríguez-Iturbe, Correlation links between normal and lognormal processes, *Water Resour. Res.*, *10*, 689–693, 1974*a*.

Mejia, J. M., and I. Rodríguez-Iturbe, On the synthesis of random field sampling from the spectrum: An application to the generation of hydrologic spatial processes, *Water Resour. Res.*, *10*, 705–711, 1974*b*.

Montroll, E. W., and M. F. Schlesinger, On $1/f$ noise and other distributions with long tails, *Proc. Natl. Acad. Sci. USA*, *79*, 3380–3383, 1982.

Orszag, S. A., Lectures on the statistical theory of turbulence, in *Fluid Dynamics*, edited by R. Balian and J.-L. Peube, pp. 235–374, Gordon and Breach, New York, 1977.

Schertzer, D., and S. Lovejoy, Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, *J. Geophys. Res.*, this issue.

Slepian, D., On the symmetrized Kronecker power of a matrix and extensions of Mehler's formula for Hermite polynomials, *SIAM J. Math. Anal.*, *3*, 606–616, 1972.

Waymire, E., V. K. Gupta, and I. Rodriguez-Iturbe, A spectral theory of rainfall intensity at the meso-β scale, *Water Resour. Res.*, *20*, 1453–1465, 1984.

Zawadzki, I. I., Statistical properties of precipitating patterns, *J. Appl. Meteorol.*, *12*, 459–472, 1973.

T. L. Bell, Laboratory for Atmospheres, Code 613, NASA Goddard Space Flight Center, Greenbelt, MD 20771.

(Received July 29, 1986;
revised March 23, 1987;
accepted March 27, 1987.)