WATER RESOURCES RESEARCH, VOL. 23, NO. 10, PAGES 1893–1900, OCTOBER 1987

Statistical Modeling of Space-Time Rainfall Using Radar and Rain Gage Observations

JAMES A. SMITH

Interstate Commission on the Potomac River Basin, Rockville, Maryland

WITOLD F. KRAJEWSKI

Hydrologic Research Laboratory, National Weather Service, Silver Spring, Maryland

A statistical framework for modeling space-time rainfall using radar and rain gage observations is developed. Three principal tasks are involved in implementing our statistical model. These tasks are referred to as sampling (that is, characterization of the error structure of radar and rain gage measurements of rainfall, modeling), (that is, specification of a stochastic model for space-time rainfall), and parameter estimation. It is emphasized that sampling, modeling, and parameter estimation are interrelated and equally important tasks. Our statistical model is applied to daily rainfall fields in the tropical Atlantic region covered by the GATE experiment.

1. INTRODUCTION

The practical need for space-time rainfall models is compelling. Rainfall is not measured continuously in space and time. Models are necessary as an adjunct to measurements. Two broad classes of models have been prominent: these are Gaussian random field models and their variants [see *Bras and Rodriguez-Iturbe*, 1985] and cluster models, introduced by *LeCam* [1961]. It is the latter class of models that we focus on in this paper. *Gupta and Waymire* [1979] (see also *Amorocho and Wu* [1975], *Waymire et al.* [1984], and *Rodriguez-Iturbe et al.* [1986]) have combined recent developments in meteorology with the LeCam modeling framework to produce very sophisticated models of space-time rainfall. It is hoped that cluster models will provide valuable tools, in conjunction with measurements of rainfall, for inferring properties of space-time rainfall. This line of research has commenced with the work of *Eagleson* [1984] and *Valdes et al.* [1985].

For the practical utilization of space-time rainfall models, we must have accurate estimates of model parameters. Development of procedurês to estimate parameters of space-time rainfall models, however, has not kept pace with development of models [*AGU Committee on Precipitation*, 1984; *Rodriguez-Iturbe*, 1986; *Georgakakos and Kavvas*, 1987]. *Smith and Karr* [1985] consider parameter estimation for cluster models using data sets consisting of rain gage observations. They conclude that severe limitations on model structure are imposed by identifiability problems resulting from the limited spatial information that can be obtained from operational rain gage networks. The identifiability problems described by Smith and Karr suggest the need for additional spatial information such as provided by radar or satellite imagery.

In this paper we present a statistical model of time-integrated rainfall fields which accommodates data from multiple sensors. Our statistical model consists of a random field $\{\xi_i(x),\ x \in \mathbf{R}^2,\ i = 1, 2, \cdots\}$, where $\xi_i(x)$ represents accumulated rainfall during time period i at spatial location x, and a data set H_n representing observations of the random field ξ and exogenous processes related to ξ available at time n. The form of our statistical model dictates three tasks. We must specify the relationshp between measurements of rainfall fields and the actual values of rainfall. This task, which we refer to as sampling, is the topic of section 2. We must also specify a probability model for rainfall fields. The model, which is developed in section 3, falls within the category of cluster models. Finally, we must develop a parameter estimation procedure for determining parameter values of the probability model. The estimation procedure presented in section 4 is a method of moments procedure. It is emphasized in the paper that modeling, sampling, and parameter estimation are equally important and interrelated tasks.

We develop a statistical model for which the data set consists of time-integrated radar observations from a single radar and time-integrated rain gage observations from a network of rain gages. The sampling model we present in section 2 is motivated by the assumptions that the strength of rain gage data is accuracy of time-integrated totals, while the strength of radar, for purposes of parameter estimation, is its ability to accurately delineate regions receiving rainfall from regions receiving no rainfall.

Smith and Karr [1985] illustrate that for estimating parameters of cluster models, maximum likelihood methods are tractable only for simple models and simple sampling situations. Consequently, reliance must generally be placed on alternative estimation procedures (and especially method of moments due to its broad applicability). An attractive feature of the method of moments estimation procedure is that a simple statistic (equation (42)) can be precomputed from the radar and rain gage observations to indicate whether feasible estimators exist.

In section 5 we apply our statistical model to the Atlantic tropical region covered by the GATE experiment [*Hudlow and Patterson*, 1979]. This experiment provides one of the best radar rainfall data sets available. Our initial result, using the feasibility statistic alluded to in the previous paragraph, is that parameter estimators cannot be obtained for the model. Further analysis using the feasibility statistic indicates that low-intensity "background" rainfall is the principal reason that the

Paper number 7W4826.
0043-1397/87/007W-4826$05.00

feasibility criterion is not satisfied. Indeed, if low-intensity rainfall is removed, we show that the feasibility criterion is satisfied and physically realistic parameter estimates are obtained.

2. SAMPLING

Parameter estimation for the rainfall model $\{\xi_i(x)\}$ will be based on observations $\{Z_i(x),\ x \in \Omega\}$ from a single radar covering a region $\Omega \subset \mathbf{R}^2$ and observations $\{\xi_i(x_j);\ j = 1, \cdots, k\}$ from k rain gages with locations $x_1, \cdots, x_k \in \Omega$. The data set at time n is thus

$$H_n = \{Z_i(x), \xi_i(x_j);\ i \leq n, j = 1, \cdots, k, x \in \Omega\} \tag{1}$$

Three assumptions contained in (1) warrant particular mention.

1. Observations from radar and rain gages are time-integrated values covering identical time intervals. Thus $\xi_i(x)$ and $Z_i(x)$ represent time-integrated values for the ith time interval of length Δt. In the example of section 5, time intervals are of length 1 day.

2. Radar observations are continuous in space. Thus for every point x in the region Ω covered by the radar, we have a value of radar reflectivity. In practice, some degree of spatial averaging will be present in any set of radar observations.

3. Time-integrated rain gage observations are error-free. Note that the rain gage observation for the jth gage is $\xi_i(x_j)$, the actual value of the rainfall field at location x_j.

Our data model is not complete without specifying the relationship between radar reflectivity $\{Z_i\}$ and the rainfall field $\{\xi_i\}$. The traditional approach of radar meteorology [*Battan*, 1973] begins with a "Z-R relationship" of the form

$$Z_i(x) = a\xi_i(x)^b \varepsilon_i(x) \tag{2}$$

where a and b are positive constants, and $\{\varepsilon_i(x),\ x \in \Omega\}$ is a nonnegative error field (the term Z-R relationship was adopted in the radar literature due to the convention of using the symbol Z for radar reflectivity and R for rainfall rate).

Wilson and Brandes [1979] review research on error structure of Z-R relationships. They note that several features are common to most studies. These include the following.

1. Radar error increases with distance from the radar. This suggests that the variance of $\varepsilon_i(x)$ is an increasing function of the distance from x to the radar.

2. Radar error is spatially correlated.

3. Radar error depends on the type of rainfall. Error structure for convective storms, for example, would differ from error structure for cyclonic storms.

A serious obstacle to using radar data for inference problems concerning the rainfall field $\{\xi_i\}$ is that (2) specifies that radar observations $\{Z_i\}$ depend on the rainfall field $\{\xi_i\}$ through parameters of the Z-R relationship (a, b, and parameters of the error field $\{\varepsilon_i\}$). From the results of *Wilson and Brandes* [1979] we can expect that accurate specification of all error field parameters will be especially difficult.

Rather than assume that parameters of the Z-R relationship are known or attempt to simultaneously estimate parameters of the Z-R relationship and rainfall model, we assume that radar can accurately distinguish regions receiving rainfall from regions receiving none, that is,

$$\xi_i(x) = 0 \qquad \text{if and only if } Z_i(x) = 0 \tag{3}$$

This assumption has been used by a number of authors in estimating storm wetted area characteristics (see, for example, *Smith* [1977], *Doneaud et al.* [1981], and *Lovejoy* [1982]).

We illustrate in section 4 that parameters of our rainfall model can be estimated from rain gage data and "0-1 mosaics" of the radar field

$$\tilde{Z}_i(x) = 1(Z_i(x) > 0) \tag{4}$$

($\tilde{Z}_i(x)$ equals 1 if $Z_i(x)$ is positive and 0 otherwise) so that it is, indeed, unnecessary to assume that parameters of the radar equation are known or to simultaneously estimate Z-R parameters and rainfall model parameters. We will denote by

$$\tilde{H}_n = \{\tilde{Z}_i(x), \xi_i(x_j);\ i \leq n, j = 1, \cdots, k, x \in \Omega\} \tag{5}$$

the reduced data set consisting of rain gage observations and 0-1 mosaics of radar fields. An attractive aspect of using radar data in the form of 0-1 mosaics is that it potentially allows the use of long records of analog radar rainfall data which, for the United States, can be obtained from the National Climatological Data Center.

3. RAINFALL MODEL

In this section we present a model for time-integrated rainfall fields $\{\xi_i(x)\}$. Temporal evolution of the model is governed by a Markov chain $\{Y_i\}$ which specifies the sequence of wet-dry periods (period i is wet if rainfall occurs anywhere in the region Ω). Structure of the rainfall field during a wet period is illustrated in Figure 1. Circular "rain cells" are organized into ellipsoidal "rainbands" which are randomly distributed in the plane. Rainfall intensity is assumed constant over a cell but varies randomly from cell to cell.

We use the terms rain cell and rainband loosely to indicate spatial scales of aggregation within the model. The smallest spatial scale in our model is represented by a cell. The largest spatial scale is the storm scale which is assumed to be larger than the region Ω. It is likely that the spatial scales we associate with rain cells and rainbands in our model will depend on the model time step [see *Rodriguez-Iturbe*, 1986].

The Markov chain model $\{Y_i\}$ of wet-dry periods contains two parameters: the probability of transition from a dry period to a dry period

$$q_0 = P\{Y_{i+1} = 0 \mid Y_i = 0\} = 1 - P\{Y_{i+1} = 1 \mid Y_i = 0\} \tag{6}$$

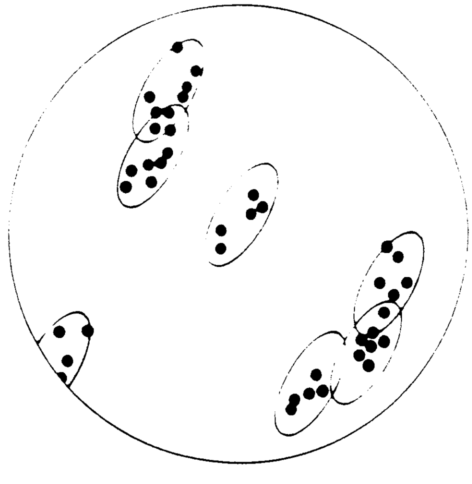

Fig. 1. Structure of the rainfall field during a wet period.

and the probability of transition from a wet period to a wet period

$$q_1 = P\{Y_{i+1} = 1 | Y_i = 1\} = 1 - P\{Y_{i+1} = 0 | Y_i = 1\} \quad (7)$$

These determine the limit probabilities

$$p_0 = \lim_{i \to \infty} P\{Y_i = 0\} \quad (8)$$

$$p_1 = \lim_{i \to \infty} P\{Y_i = 1\} = 1 - p_0 \quad (9)$$

In particular,

$$p_1 = \frac{1 - q_0}{2 - q_0 - q_1} \quad (10)$$

We assume that rain cells are circular and have fixed radius r, that is, cell radius does not vary from cell to cell or period to period. We choose to treat cell radius r in our model as not only fixed but also as known a priori; we will not try to estimate cell radius along with other model parameters. Specification of cell radius is discussed further in section 5.

The rainfall field $\{\xi_i(x)\}$ can be represented in terms of a marked point process $\{L_j^i, U_j^i\}$, where L_j^i is the spatial location of the center of the jth cell during period i and U_j^i is the storm depth (in millimeters) of the jth cell. (See *Karr* [1986] for background material on marked point processes.) We have

$$\xi_i(x) = Y_i \sum_{j=1}^{\infty} U_j^i 1(\|L_j^i - x\| \le r) \quad (11)$$

where $1(\|L_j^i - x\| \le r)$ equals 1 if the jth cell is located within distance r of x, and 0 otherwise. Note that $\xi_i(x) = 0$ for all x in Ω if $Y_i = 0$; dry periods are periods with no rainfall anywhere in Ω.

The marked point process $\{L_j, U_j\}$ can be described as follows. The locations L_j of cell centers are restricted to lie within rainbands which are modeled as ellipses (see Figure 1). Within an ellipse, locations of rain cells constitute a spatial Poisson process with rate γ (in cells per square kilometer). The assumption that rainbands are randomly located in the plane means that centroids of the ellipses constitute a spatial Poisson process with rate λ (in rainbands per square kilometer).

Geometry of rainbands is specified by three parameters, radius of the major axis a_1, (in kilometers), radius of the minor axis $a_1 a_2$ (with $0 < a_2 \le 1$), and orientation of the major axis from north to south θ_0 (with $-90° \le \theta_0 \le 90°$). The area of a rainband is $a_2 \pi a_1^2$; the mean number of rain cells per rainband is $\gamma a_2 \pi a_1^2$.

Storm depths U_j for individual rain cells are assumed to be independent and identically exponentially distributed with parameter β. The average storm depth for an individual rain cell is thus β^{-1} (in millimeters).

The probability law of a random field can be specified by its Laplace functional

$$L_{\xi_i}(f) = E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\,dx\right\}\right]$$

where f: $\mathbf{R}^2 \to \mathbf{R}_+$. The Laplace functional not only uniquely determines the probability law of ξ; it also provides distributional properties needed for parameter estimation in section 4. (See *Karr* [1986] for an expository treatment of Laplace functionals.) The following theorem contains the major distributional result for our rainfall model.

Theorem 1. The Laplace functional of $\{\xi_i\}$ is given by

$$L_{\xi_i}(f) = p_1 \exp\left\{-\lambda \int_{\mathbf{R}^2} (1 - e^{-H(z)})\,dz\right\} + (1 - p_1) \quad (12)$$

where

$$H(z) = \int_{\mathbf{R}^2} \left[\frac{\int_{\mathbf{R}^2} f(x)1(\|y - x\| \le r)\,dx}{\beta + \int_{\mathbf{R}^2} f(x)1(\|y - x\| \le r)\,dx}\right] g(z - y)\,dy \quad (13)$$

$$g(z) = \gamma \qquad z \in G$$

$$g(z) = 0 \qquad \text{otherwise}$$

and G is the region covered by an ellipse with parameters a_1, a_2, and θ_0, whose centroid is located at the origin. The proof is given in the appendix.

Using the preceding result, it is straightforward to obtain the regular Laplace transform of $\xi_i(x)$. We have

$$L_x(\alpha) = E[\exp\{-\alpha\xi_i(x)\}] = p_1$$
$$\cdot \exp\left\{-\lambda \int_{\mathbf{R}^2} [1 - e^{(-\alpha)(\alpha+\beta)\int_{R_x} g(z-y)\,dy}]\,dz\right\} + (1 - p_1) \quad (14)$$

where

$$R_x = \{y \in \mathbf{R}^2 : \|y - x\| \le r\} \quad (15)$$

is the set of points within distance r of x. We now introduce the following approximation:

$$\int_{R_x} g(z - y)\,dy \simeq \pi r^2 g(z - x) \quad (16)$$

The approximation, (16), is an equality unless x is near the boundary of the ellipse specified by g. If x is further than distance r from the boundary of the ellipse (either inside or outside), then (16) is an equality. For x within the ellipse and within r of the boundary, the right side of (16) is larger than the left. If x is outside of the ellipse and within r of the boundary, the right side of (18) is smaller than the left. Note in particular that if $\pi r^2 \ll a_2 \pi a_1^2$, that is, if rain cells are small relative to rainbands, then the error introduced by (16) is very small.

Using (12) and (16), the following distributional results for ξ_i can be obtained:

$$E[\xi_i(x)|Y_i = 1] = \lambda\beta^{-1}\pi r^2 \gamma a_2 \pi a_1^2 \quad (17)$$

$$\text{Var}\,[\xi_i(x)|Y_i = 1] = E[\xi_i(x)|Y_i = 1][\beta^{-1}(2 + \gamma\pi r^2)] \quad (18)$$

$$P\{\xi_i(x) = 0 | Y_i = 1\} = \exp\,[-\lambda a_2 \pi a_1^2 (1 - e^{-\gamma\pi r^2})] \quad (19)$$

$$P\{\xi_i(x_1) = 0, \xi_i(x_2) = 0 | Y_i = 1\} =$$
$$\exp\,[-\lambda(2a_2\pi a_1^2 - A_{x_1x_2})(1 - e^{-\gamma\pi r^2})] \quad (20)$$

where $A_{x_1x_2}$ is the area of overlap of two ellipses (with parameters a_1, a_2, and θ_0) whose centroids are at x_1 and x_2. We present distributional properties conditioned on $Y_i = 1$ because these are the results needed for the estimation procedure of section 4.

4. PARAMETER ESTIMATION

In this section we develop method of moments estimators for parameters of the rainfall model presented in section 3. Recall from section 2 that our data set for parameter estimation is

$$\vec{H}_n = \{\xi_i(x_j), \tilde{Z}_i(x);\ i \le n, j = 1, \cdots, k, x \in \Omega\}$$

where $\tilde{Z}_i(x) = 1$ if radar indicates positive rainfall at location x and $\tilde{Z}_i(x) = 0$ if radar indicates no rain at x. We will use the random variable

$$\tilde{Z}_i = 1\left(\int_\Omega \tilde{Z}_i(x)\, dx > 0\right) \tag{21}$$

to indicate whether rainfall is observed anywhere in our region Ω by radar ($\tilde{Z}_i = 1$) or whether radar shows the entire region to be dry ($\tilde{Z}_i = 0$).

One final assumption on our data set will prove quite useful; we assume that Ω is large enough that the probability of no rain cells in Ω during a wet period is very small. This assumtion implies that $\tilde{Z}_i = Y_i$ with high probability.

It follows that the estimation problem can be separated into two parts: estimation of temporal parameters associated with the Markov chain model of wet-dry sequences and estimation of spatial parameters associated with the random field model. Note that more complex models of wet-dry sequences (see, for example, *Foufoula-Georgiou and Lettenmaier* [1987] and *Smith* [1987]) could be used without affecting the spatial component of the model.

Estimation of temporal parameters of the Markov chain model is straightforward under the assumption that $\tilde{Z}_i = Y_i$. We estimate q_0 by

$$\hat{q}_0 = \sum_{i=2}^{n} (1 - \tilde{Z}_i)(1 - \tilde{Z}_{i-1}) \left(\sum_{i=1}^{n-1} (1 - \tilde{Z}_i)\right)^{-1} \tag{22}$$

and q_1 by

$$\hat{q}_1 = \sum_{i=2}^{n} \tilde{Z}_i \tilde{Z}_{i-1} \left(\sum_{i=1}^{n-1} \tilde{Z}_i\right)^{-1} \tag{23}$$

In words, $\hat{q}_1$ is the fraction of wet periods that are followed by wet periods and $\hat{q}_0$ is the fraction of dry periods followed by dry periods.

For estimation of spatial parameters we use a method of moments procedure which employs the following statistics.

1. An estimator of $E[\xi_i(x)|Y_i = 1]$ is given by

$$\hat{\mu} = \left(\sum_{i=1}^{n} \tilde{Z}_i\right)^{-1} \sum_{i=1}^{n} \tilde{Z}_i \left[k^{-1} \sum_{j=1}^{k} \xi_i(x_j)\right] \tag{24}$$

2. An estimator of Var $(\xi_i(x)| Y_i = 1)$ is given by

$$\hat{\sigma}^2 = \left(\sum_{i=1}^{n} \tilde{Z}_i\right)^{-1} \sum_{i=1}^{n} \tilde{Z}_i \left[k^{-1} \sum_{j=1}^{k} (\xi_i(x_j) - \hat{\mu})^2\right] \tag{25}$$

3. An estimator of $P\{\xi_i(x) = 0| Y_i = 1\}$ is given by

$$\hat{p} = \left(\sum_{i=1}^{n} \tilde{Z}_i\right)^{-1} \sum_{i=1}^{n} \tilde{Z}_i \left[|\Omega|^{-1} \int_\Omega (1 - \tilde{Z}_i(x))\, dx\right] \tag{26}$$

where $|\Omega|$ is the area of Ω.

4. An estimator of $p(\theta, s) = P\{\xi_i(x_1) = 0, \xi_i(x_2) = 0| Y_i = 1\}$, (where $x_1 - x_2 = (\theta, s)$, that is, x_1 and x_2 are separated by distance s and the line segment between them is oriented θ degrees from north to south) is given by

$$\hat{p}(\theta, s) = \left(\sum_{i=1}^{n} \tilde{Z}_i\right)^{-1} \sum_{i=1}^{n} \cdot \tilde{Z}_i \left[|\tilde{\Omega}|^{-1} \int_{\substack{\tilde{\Omega} \\ x - y = (\theta, s)}} (1 - \tilde{Z}_i(x))(1 - \tilde{Z}_i(y))\, dx\right] \tag{27}$$

where $\tilde{\Omega}$ is the set of points x in Ω such that the point y which is located at distance s and orientation θ from x is also in Ω. The area of $\tilde{\Omega}$, $|\tilde{\Omega}|$, will, of course, be smaller than the area of Ω.

Note that the first two statistics require both radar and rain gage observations; the latter two statistics use only radar data. To estimate the conditional probability of no rainfall at a point (equation (26)), given that the period is a wet period, we simply average the dry area values for all storm days. Similarly, to estimate the joint conditional probability of no rainfall at two points (equation (27)) separated by distance s and angle θ, we average the product of observations (that are separated by distance s and angle θ) both spatially and over time, for all storm days.

Estimation of θ_0 is based on the observation that

$$p(\theta_0, s) \ge p(\theta, s) \qquad \forall\ s \text{ and } \theta$$

$$p(\theta_0, s) > p(\theta, s) \qquad \theta \ne \theta_0 \qquad s < a_1 a_2$$

which is obtained from (20). We choose $\hat{\theta}_0$ as the solution to the optimization problem

$$\max \left\{H(\theta) = \int_0^{t_0} \hat{p}(\theta, s)\, ds;\ -90^\circ \le \theta \le 90^\circ\right\}$$

where t_0 is chosen to ensure satisfactory sampling properties of H. Our estimate of ellipse orientation is thus the direction (from the radar) along which the largest joint probability of zero rainfall occurs.

We now define

$$\hat{p}_1(s) = \hat{p}(\hat{\theta}_0, s) \tag{28}$$

$$\hat{p}_2(s) = \hat{p}(\hat{\theta}_0 - 90, s) \tag{29}$$

Thus $\hat{p}_1(s)$ is an estimator of the probability of zero rainfall at two points separated by distance s oriented along the major axis of a rainband while $\hat{p}_2(s)$ is analogous for minor axis. Our moment equations can now be expressed as follows using (17)–(20):

$$\hat{\mu} = \lambda\beta^{-1}\gamma\pi r^2 a_2 \pi a_1^{\ 2} \tag{30}$$

$$\hat{\sigma}^2 = \hat{\mu}[\beta^{-1}(2 + \gamma\pi r^2)] \tag{31}$$

$$-\log(\hat{p}) = \lambda a_2 \pi a_1^{\ 2}(1 - e^{-\gamma\pi r^2}) \tag{32}$$

$$-\log(\hat{p}_1(s)) = \lambda[2a_2\pi a_1^{\ 2} - A_1(s)](1 - e^{-\gamma\pi r^2}) \tag{33}$$

$$-\log(\hat{p}_2(s)) = \lambda[2a_2\pi a_1^{\ 2} - A_2(s)](1 - e^{-\gamma\pi r^2}) \tag{34}$$

where $A_1(s)$ is the area of overlap of two ellipses oriented along their major axes with centroids separated by distance s; the definition for $A_2(s)$ is analogous with "minor axis" replacing "major axis." For orientation along major axes, we have

$$A_1(s) = a_2\pi a_1^{\ 2} - s a_2(a_1^{\ 2} - 0.25 s^2)^{1/2} - 2a_1^{\ 2} a_2 \arcsin(0.5 s a_1^{\ -1}) \tag{35}$$

Combining (32) and (33) and (32) and (34) we obtain

$$\frac{\log(\hat{p}_1(s))}{\log(\hat{p})} = \frac{2a_2\pi a_1^{\,2} - A_1(s)}{a_2\pi a_1^{\,2}} \tag{36}$$

$$\frac{\log(\hat{p}_2(s))}{\log(\hat{p})} = \frac{2a_2\pi a_1^{\,2} - A_2(s)}{a_2\pi a_1^{\,2}} \tag{37}$$

Note that the right-hand sides of both (36) and (37) depend only on a_1 and a_2. We choose $\hat{a}_1$ and $\hat{a}_2$ as solution to the optimization problem

$$\min_{a_1,a_2} \hat{H}(a_1, a_2) = \int_0^{t_0} \left[\frac{\log(\hat{p}_1(s))}{\log(\hat{p})} - \left(\frac{2a_2\pi a_1^{\,2} - A_1(s)}{a_2\pi a_1^{\,2}}\right)\right]^2 + \left[\frac{\log(\hat{p}_2(s))}{\log(\hat{p})} - \left(\frac{2a_2\pi a_1^{\,2} - A_2(s)}{a_2\pi a_1^{\,2}}\right)\right]^2 ds \tag{38}$$

such that $a_1 > 0$ and $0 < a_2 \leq 1$. As before, t_0 is chosen to ensure satisfactory sampling properties of $\hat{H}$.

Our remaining moment equations can be rewritten as follows:

$$\beta^{-1} = (\hat{\sigma}^2/\hat{\mu})(2 + \gamma\pi r^2)^{-1} \tag{39}$$

$$\lambda = -\log(\hat{p})(\hat{a}_2\pi\hat{a}_1^{\,2})^{-1}(1 - e^{-\gamma\pi r^2})^{-1} \tag{40}$$

$$\frac{-\log(\hat{p})\hat{\sigma}^2}{\hat{\mu}^2} = \frac{2 + \gamma\pi r^2}{\gamma\pi r^2}\,[1 - e^{-\gamma\pi r^2}] = f(\gamma) \tag{41}$$

Note that

$$f(0) = 2$$

$$\lim_{\gamma \to \infty} f(\gamma) = 1$$

$$f'(\gamma) < 0 \qquad \text{for } \gamma > 0$$

implying existence of $f^{-1}: (1, 2] \to [0, \infty)$. If

$$1 < \frac{-\log(\hat{p})\hat{\sigma}^2}{\hat{\mu}^2} < 2 \tag{42}$$

we choose $\hat{\gamma}$ as the unique solution to (41), that is,

$$\hat{\gamma} = f^{-1}\left(\frac{-\log(\hat{p})\hat{\sigma}^2}{\hat{\mu}^2}\right) \tag{43}$$

Finally, we take

$$\hat{\beta}^{-1} = (\hat{\sigma}^2/\hat{\mu})(2 + \hat{\gamma}\pi r^2)^{-1} \tag{44}$$

$$\hat{\lambda} = -\log(\hat{p})(\hat{a}_2\pi\hat{a}_1^{\,2})^{-1}(1 - e^{-\hat{\gamma}\pi r^2})^{-1} \tag{45}$$

Condition (42) is quite useful in practice. As a first step in carrying out the estimation procedure one computes the statistic in (42). If it is between 1 and 2, existence of parameter estimates is guaranteed. If the statistic does not fall in this range, estimates do not exist. In the latter case, one concludes that either the model is inappropriate for the data set or that the data set is too short. Unfortunately, there is no clear way of distinguishing the two cases.

It can be shown that our method of moment estimators possess attractive large sample properties typically associated with maximum likelihood estimators (see W. F. Krajewski and J. A. Smith, unpublished manuscript, 1987). Specifically, our estimators are asymptotically normal and consistent in mean square error.

The proof of asymptotic normality is straightforward and can be imitated in many situations in which method of moments estimators are used. There are three basic steps: one shows that sample statistics (equations (24)–(27)) are asymptotically normal using a standard central limit theorem; (2) one shows that parameter estimators are "reasonable" functions of sample statistics; and (3) one invokes an appropriate theorem (see, for example, *Serfling* [1980]) stating that reasonable functions of asymptotically normal statistics are themselves asymptotically normal.

5. RESULTS OF MODEL APPLICATION

In this section we describe application of our statistical model to daily rainfall in the tropical Atlantic. We choose this area due to availability of radar rainfall data from the GARP Atlantic Tropical Experiment (GATE). The GATE experiment provides a very high quality radar rainfall data set which covers a region of approximately 120,000 km^2 (for detailed discussion of the GATE data set, see *Hudlow and Patterson* [1979]). Modeling based on the GATE data set has played, and continues to play, an important role in research on the hydrology, meteorology, and oceanography of tropical oceanic regions (see, for example, *Bell* [1987]).

Because there were only a few rain gages used during the GATE project, we simulated a sampling mechanism for the rain gage network. This approach is justified in the case of GATE radar data due to its high quality and good agreement with the rain gage data actually used [see *Hudlow et al.* 1979]. The procedure we used to generate rain gage observations from the radar rainfall GATE data accounts for different sampling properties of the two sensors and is described in the work by *Krajewski* [1987].

The estimation procedure of section 4 was applied using daily data for the 57 days of three phases of GATE. Daily radar fields were constructed by aggregating the hourly radar fields (which are summarized in the work by *Hudlow and Patterson* [1979]). Using a simulated rain gage network of 400 gages (for a gage density of approximately 1 gage/300 km^2) we first computed the feasibility statistic of (42) obtaining a value of 4.34. This result implies that estimators of model parameters cannot be computed.

From (42) we can loosely conclude that mean rainfall and the probability of no rainfall are too small and the variance of rainfall is too large. It is argued below that the crux of the problem is that the probability of no rainfall during wet periods is too low. In effect, the model is not capable of handling low-intensity background rainfall. To arrive at this conclusion, we created censored rainfall fields in the following fashion. Daily radar fields were constructed after first removing from the hourly fields values less than a specified threshold. In Table 1 we show the estimated mean, variance, probability of no rainfall, and feasibility statistic for the original data set and censored data sets with thresholds of 0.5 and 1.0 mm/hour. Note that the mean and variance for the 0.5 threshold differ from the zero threshold values by less than 12%; the probability of no rainfall with the 0.5 threshold is more than twice the zero threshold value.

Note also in Table 1 that the feasibility statistics for both the 0.5 and 1.0 mm/hour thresholds lie between 1 and 2. Clearly, the feasibility statistics for the censored rainfall fields are in the appropriate range due to the increase in the probability of no rainfall. The mean and variance are not only slowly changing with increasing threshold, they are also moving in the wrong direction. The sample mean is decreasing

TABLE 1. Statistics Computed for Various Thresholds

Statistic	Threshold, mm/hour 0.0	0.5	1.0
$\hat{\mu}$	0.46	0.40	0.33
$\hat{\sigma}^2$	0.78	0.81	0.85
$\hat{p}$	0.30	0.73	0.85
$\hat{s}$	4.34	1.61	1.28

with increasing threshold. If the other parameters remain constant, a decreasing sample mean increases the feasibility statistic. Similarly, an increasing sample variance would increase the feasibility statistic if other parameters remain constant. Table 2 contains values of the feasibility statistic for a range of rain gage numbers. These results rule out sampling properties of the rain gages as a problem source. These results strongly suggest that shortcomings of the model for the GATE region relate to low-intensity rainfall.

Using the radar fields censored at 0.5 mm hour and the parameter estimation procedure of section 4, the following parameter estimates were obtained: $\hat{\lambda} = 0.0003$ rainbands km^{-2}, $\hat{\gamma} = 0.14$ rain cells km^{-2}, $\hat{\beta}^{-1} = 0.63$ mm, $\hat{a}_1 = 20.4$ km, and $\hat{a}_2 = 1.0$. The value of rain cell radius was taken to be 2 km. Note in (17)–(20) that rain cell radius appears only through terms involving $\gamma\pi r^2$. Thus if we change r, the only change in model parameters involves γ. If, for example, we use a rain cell radius of 1 km, the rain cell intensity γ increase to 0.54. Using a value of 2 km gives us rain cells that are approximately the size of our radar pixels.

It is noteworthy that the estimated model has rain bands that are approximately circular in shape (that is, $\hat{a}_2 = 1.0$); this is not overly surprising for tropical rainfall. It would be expected that extratropical storms would show stronger anisotropic organization.

From our parameter estimates, we conclude that during a wet day our 120,000 km^2 region contains, on average, 36 rainbands. An individual rain band, which has an area of approximately 1300 km^2, contains, on average, 182 rain cells. Average rainfall intensity for a single rain cell is 0.63 mm.

In our application we have identified one of the possible limitations of the model: its inability to deal with low-intensity rain hiding more apparent structure of a rainfall field. Other limitations such as those attributable to small sample properties of parameter estimators, effects of rain gage density, and the structure of the rainall process itself could be examined via a Monte Carlo study.

TABLE 2. Feasibility Statistic Computed for Various Network Densities

Number of Gages	Threshold, mm hour 0.00	0.50	1.00
25	4.89	1.72	1.35
50	4.94	1.78	1.38
100	4.59	1.65	1.29
200	4.28	1.57	1.27
400	4.34	1.61	1.28
800	4.43	1.63	1.29

6. CONCLUSIONS

In this paper we have presented a statistical framework for multisensor rainfall modeling; three principal tasks are involved. In section 2 we present our sampling model which characterizes the error of radar and rain gage measurements of rainfall. Our sampling model is based on the assumptions that, for purposes of parameter estimation, the strength of rain gage data is accuracy of time-integrated obervations while the strength of radar is its ability to "see" areal extent of rainfall fields. In section 3 we develop a cluster model for time-integrated rainfall fields. The major distributional result we obtained for our model, a representation for its Laplace functional, is given in theorem 1. From this result, we are able to compute the theoretical moments necessary for parameter estimation. Method of moments estimators for model parameters are derived in section 4. The estimators require rain gage data and 0-1 mosaics of radar fields. An especially attractive feature of the estimation procedure is that a simple statistic (equation (42)) can be precomputed to determine whether feasible parameter estimates exist.

It has been shown that inclusion of radar rainfall data, even in the simple form of 0-1 mosaics, improves the model identification problems discussed by *Smith and Karr* [1985]. By using radar data in conjunction with rain gage data, we can consider more complex rainfall models than if rain gage data alone are used. If, however, we wish to consider more complex rainfall models than the one presented in section 3, several problems may arise. It may not be possible to obtain the information necessary for parameter estimation from 0-1 mosaics of radar fields, requiring direct use of high-resolution digitized radar fields. This, in turn, would likely require stronger assumptions on radar error structure than were made in section 2. A second problem that may arise is loss of computational tractability. In choosing a more complex model, we may lose the capability to compute distributional properties that are necessary for parameter estimation. Ultimately, we must balance the validity of rainfall model assumptions against the validity of sampling assumptions and analytical tractability.

In section 5 we apply our statistical model to the tropical Atlantic region covered by the GATE experiment. Results of this section suggest that our rainfall model is suitable for important components of tropical Atlantic rainfall. The model is not capable, however, of representing low-intensity components of the rainfall process. The results of section 5 reinforce our conviction that progress in rainfall analysis is best achieved by a coordinated and balanced treatment of modeling, sampling, and parameter estimation.

APPENDIX

The proof of theorem 1 is sketched below. First, we note that an alternative representation to (11) for the random field ξ is

$$\xi_i(x) = Y_i \int_{\mathbf{R}^2 \times \mathbf{R}_+} u 1(|y - x| \le r)\, dM_i(y, u) \qquad \text{(A1)}$$

where M_i is a point process on $\mathbf{R}^2 \times \mathbf{R}_+$. For $A \subset \mathbf{R}^2$ and $B \subset \mathbf{R}_+$, $M_i(A, B)$ is the number of rain cells whose center is located in A and whose storm depth is contained in B.

Computational tractability of ξ_i follows from the fact that

the point process M_i in (A1) can be represented as a Cox process on $\mathbf{R}^2 \times \mathbf{R}_+$. A Cox process is a point process which can be interpreted as a Poisson process with a randomly varying rate of occurrence [see *Karr*, 1986]. The randomly varying rate of occurrence will be termed the "directing process." The directing process of M_i is given by

$$\Lambda_i(x, u) = \tilde{\Lambda}_i(x)h(u) \qquad x \in \mathbf{R}^2 \quad u \in \mathbf{R}_+ \tag{A2}$$

where h is the exponential density function with parameter β and

$$\tilde{\Lambda}_i(x) = \int_{\mathbf{R}^2} g(x-y)\, dN_i(y) \tag{A3}$$

N_i is a Poisson process on $\mathbf{R}^2$ with intensity λ, and g is given by (13). In this formulation, N_i is the point process of rain band centers and $g(x)$ is the rate of occurrence of rain cells at x associated with a rainband centered at the origin.

The following lemma [*Karr*, 1986] is needed for computing the Laplace functional of $\{\xi_i\}$.

Lemma

Let M be a Cox process on a Euclidean space E with directing process Λ. Then

$$E\left[\exp\left\{-\int_E f(x)\, dM(x)\right\}\right] = E\left[\exp\left\{-\int_E (1-e^{-f(x)})\Lambda(x)\, dx\right\}\right] \tag{A4}$$

Proof of Theorem 1

$$L_\xi(f) = E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\right] \tag{A5}$$

$$L_\xi(f) = E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y_i = 1\right]P\{Y_i = 1\} + E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y_i = 0\right]P\{Y_i = 0\} \tag{A6}$$

$$L_\xi(f) = E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y_i = 1\right]p_1 + (1-p_1) \tag{A7}$$

$$E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y_i = 1\right] = E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\left[\int_{\mathbf{R}^2\times\mathbf{R}_+} u1(\|y-x\| \le r)\, dM_i(y,u)\right]dx\right\}\right] \tag{A8}$$

$$E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y_i = 1\right] = E\left[\exp\left\{-\int_{\mathbf{R}^2\times\mathbf{R}_+} \cdot\left[u\int_{\mathbf{R}^2} f(x)1(\|y-x\| \le r)\, dx\right]dM_i(y,u)\right\}\right] \tag{A9}$$

$$E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y_i = 1\right] = E\left[\exp\left\{-\int_{\mathbf{R}^2\times\mathbf{R}_+} \cdot\left[1-\exp\left(-u\int_{\mathbf{R}^2} f(x)1(\|y-x\| \le r)\, dx\right)\right]\tilde{\Lambda}_i(y)h(u)\, du\, dy\right\}\right] \tag{A10}$$

$$E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y_i = 1\right] = E\left[\exp\left\{-\int_{\mathbf{R}^2}\left[\int_{\mathbf{R}_+} h(u) - h(u) \cdot \exp\left(-u\int_{\mathbf{R}^2} f(x)1(\|y-x\| \le r)\, dx\right)du\right]\tilde{\Lambda}_i(y)\, dy\right\}\right] \tag{A11}$$

$$E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y_i = 1\right] = E\left[\exp\left\{-\int_{\mathbf{R}^2} \cdot\left[1-\frac{\beta}{\beta+\int_{\mathbf{R}^2} f(x)1(\|y-x\| \le r)\, dx}\right]\tilde{\Lambda}_i(y)\, dy\right\}\right] \tag{A12}$$

Let

$$G(y) = \frac{\int_{\mathbf{R}^2} f(x)1(\|y-x\| \le r)\, dx}{\beta+\int_{\mathbf{R}^2} f(x)1(\|y-x\| \le r)\, dx} \tag{A13}$$

Then

$$E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y = 1\right] = E\left[\exp\left\{-\int_{\mathbf{R}^2} G(y)\tilde{\Lambda}_i(y)\, dy\right\}\right] \tag{A14}$$

$$E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y = 1\right] = E\left[\exp\left\{-\int_{\mathbf{R}^2} G(y)\int_{\mathbf{R}^2} g(z-y)\, dN_i(z)\, dy\right\}\right] \tag{A15}$$

$$E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y = 1\right] = E\left[\exp\left\{-\int_{\mathbf{R}^2}\left[\int_{\mathbf{R}^2} G(y)g(z-y)\, dy\right]dN_i(z)\right\}\right] \tag{A16}$$

$$E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y = 1\right] = \exp\left\{-\lambda\int_{\mathbf{R}^2}\left(1-\exp\left(-\int_{\mathbf{R}^2} G(y)g(z-y)\, dy\right)\right)dz\right\} \tag{A17}$$

$$E\left[\exp\left\{-\int_{\mathbf{R}^2} f(x)\xi_i(x)\, dx\right\}\middle| Y = 1\right] = \exp\left\{-\lambda\int_{\mathbf{R}^2} (1-e^{-H(z)})\, dz\right\} \tag{A18}$$

Key steps in the proof are (A8) and (A17), which follow from the lemma, and (A12) in which we evaluate the normal Laplace transform of an exponential random variable.

Acknowledgments. The research of J. Smith was carried out at the Centre for Mathematics and Computer Science in Amsterdam with the support of a Fulbright postdoctoral research grant. The numerical computations were performed at the Hydrologic Research Laboratory of the National Weather Service in Silver Spring, Maryland. The comments of E. Foufoula-Georgiou are gratefully acknowledged. The authors gratefully acknowledge excellent typing of the manuscript by E. Hauschildt of the Hydrologic Research Laboratory.

REFERENCES

AGU Committee on Precipitation, A new interdisciplinary focus on precipitation research, *Eos Trans. AGU*, *65*(23), 377–380, 1984.

Amorocho, J., and B. Wu, Mathematical models for the simulation of cyclonic storm sequences and precipitation fields, *J. Hydrol.*, *32*(3/4), 329–345, 1977.

Battan, L. J., *Radar Observation of the Atmosphere*, 324 pp., The University of Chicago Press, Chicago, Ill., 1973.

Bell, T. L., A space-time stochastic model of rainfall for satellite remote sensing studies, *J. Geophys. Res.*, in press, 1987.

Bras, R. L., and I. Rodriguez-Iturbe, *Random Functions and Hydrology*, 559 pp., Addison-Wesley, Reading, Mass., 1985.

Doneaud, A., P. S. Smith, A. Dennis, and S. Sengupta, A simple method for estimating convective rain volume over an area, *Water Resour. Res.*, *17*(6), 1676–1682, 1981.

Eagleson, P. S., The distribution of catchment coverage by stationary rainstorms, *Water Resour. Res.*, *20*(5), 581–590, 1984.

Foufoula-Georgiou, E., and D. Lettenmaier, A Markov renewal model for rainfall occurrences, *Water Resour. Res.*, *23*(5), 875–884, 1987.

Georgakakos, K. P., and M. L. Kavvas, Precipitation analysis, modeling, and prediction in hydrology, *Rev. Geophys.*, *25*(2), 163–178, 1987.

Gupta, V. K., and E. Waymire, A stochastic kinematic study of subsynoptic space-time rainfall, *Water Resour. Res.*, *15*(3), 637–644, 1979.

Hudlow, M. D., and V. L. Patterson, GATE Radar Rainfall Atlas, NOAA Special report, p. 155, U.S. Dep. of Commerce, Washington, D. C., 1979.

Hudlow, M. D., R. Arkell, V. Patterson, P. Pytlowany, F. Richards, and S. Goetis, Calibration and Intercomparison of the GATE C Band Radars, *NOAA Tech. Rep. EDIS 31*, U.S. Dep. of Commerce, Washington, D. C., 1979.

Karr, A. F., *Point Processes and Their Statistical Inference*, Dekker, New York, 1986.

Krajewski, W. F., Cokriging of radar and rain gage data, *J. Geophys. Res.*, in press, 1987.

LeCam, L., A stochastic description of precipitation, in *IV Berkeley Symposium on Mathematical Statistics and Probability*, vol. 3, pp. 165–186, University of California, Berkeley, 1961.

Lovejoy, S., Area-perimeter relation for rain and cloud areas, *Science*, *216*, 185–187, 1982.

Rodriguez-Iturbe, I., Scale of fluctuation of rainfall models, *Water Resour. Res.*, *22*(9), 15s–37s, 1986.

Rodriguez-Iturbe, I., D. R. Cox, and P. S. Eagleson, Spatial modeling of storm rainfall, *Proc. R. Soc. London, Ser. A*, *403*, 27–50, 1986.

Serfling, R. J., *Approximation Theorems of Mathematical Statistics*, John Wiley, New York, 1980.

Smith, D. L., An examination of probability of precipitation forecasts in light of rainfall areal coverage, *Natl. Weather Dig.*, *2*(2), 15–26, 1977.

Smith, J. A., Statistical modeling of daily rainfall occurrences, *Water Resour. Res.*, *23*(3), 885–893, 1987.

Smith, J. A., and A. F. Karr, Parameter estimation for a model of space-time rainfall, *Water Resour. Res.*, *21*(8), 1251–1257, 1985.

Valdes, J. B., I. Rodriguez-Iturbe, and V. K. Gupta, Approximations of temporal rainfall from a multidimensional model, *Water Resour. Res.*, *21*(8), 1259, 1985.

Waymire, E., C., V. K. Gupta, and I. Rodriguez-Iturbe, A spectral theory of rainfall intensity at the meso β scale, *Water Resour. Res.*, *20*(10), 1453–1465, 1984.

Wilson, J. W., and E. A. Brandes, Radar measurement of rainfall—A summary, *Bull. Am. Meteorol. Soc.*, *60*(9), 1048–1058, 1979.

W. F. Krajewski, Hydrologic Research Laboratory, National Weather Service, 8060 13th Street, Room 530, Silver Spring, MD 20910.

J. A. Smith, Interstate Commission on the Potomac River Basin, 6110 Executive Boulevard, Suite 300, Rockville, MD 20852.

(Received February 9, 1987;
revised June 22, 1987;
accepted June 22, 1987.)