WATER RESOURCES RESEARCH, VOL. 25, NO. 9, PAGES 2067–2075, SEPTEMBER 1989

Sampling Properties of Parameter Estimators for a Storm Field Rainfall Model

WITOLD F. KRAJEWSKI

Iowa Institute of Hydraulic Research, University of Iowa, Iowa City

JAMES A. SMITH

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

A statistical model of rainfall fields has been developed. The model parameters can be estimated from radar and rain gage data. The radar data are used only to estimate the spatial features of the model. The rain gage data are used to estimate the magnitude of rainfall. The parameter estimators are based on the method of moments and are shown to be consistent and asymptotically normal. To investigate the small sample properties of the estimators a Monte Carlo simulation experiment has been conducted. The results indicate that for certain combinations of the true rainfall field parameters the estimation procedure leads to biased results. The model has an attractive feature in that a simple statistic can be precomputed which indicates the feasibility of the model to represent a given rainfall data set. In case the statistic indicates infeasibility of the model it is difficult to distinguish whether the proposed model is not appropriate or the data sample is too small.

1. INTRODUCTION

Recent progress in statistical modeling of space-time rainfall brings to the forefront the problem of parameter estimation. The practical value of models can be utilized only if there is a coordinated and balanced treatment of modeling and parameter estimation. In recent years, research toward model development has significantly outpaced research on parameter estimation (for discussion, see *American Geophysical Union Committee on Precipitation* [1984], *Rodriguez-Iturbe et al.* [1986], and *Georgakakos and Kavvas* [1987]).

Traditionally, rainfall analyses have been based on rain gage data. These point measurements are generally adequate for temporal models of rainfall, but even then the situation is often complicated by the lack of scale compatibility between models and data [*Rodriguez-Iturbe et al.*, 1984; *Valdes et al.*, 1985]. Rain gage networks provide limited spatial information resulting in identifiability problems in space-time rainfall modeling [*Smith and Karr*, 1985]. To resolve spatial structure of rainfall fields, information from remote sensors such as meteorological radars or satellites is necessary. Few attempts to use radar information in conjunction with stastical models of rainfall are described in the literature; examples include the works by *Kavvas and Herd* [1985] and *Smith and Krajewski* [1987].

This paper is a continuation of the model development initiated by *Smith and Krajewski* [1987]. They have developed a statistical model of space-time rainfall for which parameter estimation is based on time-integrated radar rainfall observations from a single radar and time-integrated rain gage observations from a network of rain gages. The reason for using both types of rainfall observations stems from the fact that raingage measurements of time-integrated rainfall are more accurate than radar estimates, while radar observations are better in delineating areas receiving rainfall (for a discussion of radar rainfall estimation methods see, *Battan* [1973] and *Doviak and Zrnic* [1984]). *Smith and Krajewski* [1987] show that the parameter estimation procedure for their model, which is based on the method of moments, possesses some attractive features. Notably, a simple statistic can be precomputed from the radar and raingage data to indicate the existence of a solution to the parameter estimation problem. The estimation procedure is applied to daily rainfall fields in the tropical Atlantic region covered by the GATE experiment. It is shown that physically realistic parameter estimates are obtained using the estimation procedure.

However, the capability of the model to represent actual rainfall fields cannot be properly assessed without addressing the issue of model data requirements. Since the model parameters should be estimated from observations of a particular rainfall regime, it is important to recognize the data requirements of the model and the limitations of the proposed parameter estimators. The accuracy of the parameter estimators is affected by the estimation procedure, data measurement errors, and data sample size. Without studying all of these effects one cannot positively determine the ability of the model to represent actual rainfall fields.

The main objective of this paper is to investigate sampling properties of the parameter estimators. To facilitate the study the original model of *Smith and Krajewski* [1987] has been slightly simplified. The original anisotropic model has been modified to an isotropic version. This simplification is justified in the light of the results presented in the original paper [*Smith and Krajewski*, 1987], where the full model was applied to GATE data [*Hudlow and Patterson*, 1979] and estimated parameters suggested isotropy. Also, *Bell* [1987] used an isotropic model for GATE data. Of course, for rainfall regimes where spatial anisotropy is significant the full version of the model should be used. The simplified model is described in section 2. In section 3 the parameter estimation procedure is presented and "large sample" prop-

Paper number 89WR01188.
0043-1397/89/89WR-01188$05.00

erties of parameter estimators are derived. It is shown that the estimators are consistent and asymptotically normal. "Small sample" properties of parameter estimators are assessed through a Monte Carlo experiment. Design of the experiment and its results are the topics of section 4. A summary and conclusions are presented in section 5.

2. RAINFALL MODEL

In this section a model for storm fields $\{Y(x); x \in R\}$ is presented. The storm field is modeled as a nonnegative random field [see *Bras and Rodriguez-Iturbe*, 1985] over the region R. The random variable $Y(x)$ represents rainfall rate (or accumulated rainfall) at spatial location x. The region R will typically represent the area covered by a single sensor or group of sensors. Before presenting the model, some definitions are needed.

A nonnegative random field $\{Y(x)\}$ is "intermittent" if

$$P\{Y(x) = 0\} > 0 \qquad \forall x \in R \tag{1}$$

where P stands for probability. An intermittent random field is one that can be decomposed into a region of zero values (nonraining) and positive values. The distinction is important in this work both from a modeling perspective and in the context of parameter estimation. Radar can perform quite well in delineating regions of rainfall from regions of no rainfall.

A random field $\{Y(x)\}$ is homogeneous (or stationary) if the finite dimensional distributions of $Y(x_1 + x), \cdots, Y(x_k + x)$ are the same as those of $Y(x_1), \cdots, Y(x_k)$ for all $k, x_1, \cdots, x_k$, and x. The following distributional descriptors for intermittent, homogeneous random fields are used for parameter estimation:

$$\mu = E\{Y(x)\} \tag{2}$$

$$\sigma^2 = \mathrm{Var}\,\{Y(x)\} \tag{3}$$

$$p = P\{Y(x) = 0\} \tag{4}$$

$$q(s) = P\{Y(x_1) = 0,\ Y(x_2) = 0\} \qquad \|x_1 - x_2\| = s \tag{5}$$

The "index of variability," I, of a homogeneous, intermittent random field is defined by

$$I = -\log(p)(\sigma/\mu)^2 = -\log(p)C_v^2 \tag{6}$$

where $C_v = \sigma/\mu$ is the coefficient of variation. The index of variability can be viewed as a scaled coefficient of variation.

Circular "rain cells" of radius r are the basic building blocks of the rainfall model. The rainfall field $\{Y(x)\}$ can be represented in terms of the locations of cell centers $\{L_j\}$ and cell rainfall intensities $\{U_j\}$, as follows:

$$Y(x) = \sum_{j=1}^{\infty} 1(\|L_j - x\| < r)U_j \tag{7}$$

where $1(\|L_j - x\| < r)$ equals 1 if the jth cell is located within distance r of x, and 0, otherwise. Computational tractability of the model follows from the fact that $Y(x)$ is represented as a random sum of independent and identically distributed (iid) exponential random variables

$$Y(x) = \sum_{j=1}^{N(x)} U_{T_j} \tag{8}$$

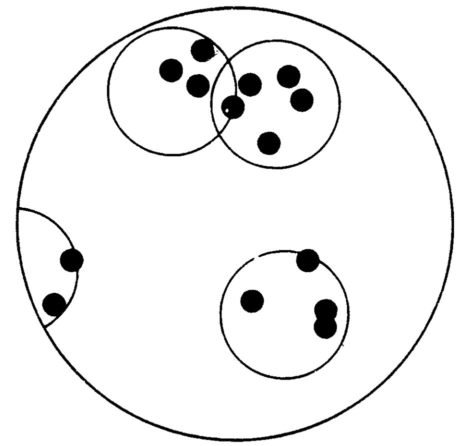

Fig. 1 Schematic representation of the model structure.

where

$$N(x) = \sum_{j=1}^{\infty} 1(\|L_j - x\| < r) \tag{9}$$

is the number of cells within distance r of the point x and T_j is the index of the jth cell which is within distance r of the point x.

The locations L_j of cell centers are restricted to lie within circular rain bands of radius a (with units in kilometers; see Figure 1). Within a rain band, locations of rain cell centers L_j constitute a spatial Poisson process with rate γ (with units in cells per square kilometer). The assumption that rain bands are randomly located in the plane means that centers of rain bands constitute a spatial Poisson process with rate λ (with units in rain bands per square kilometer). Storm depths U_j for rain cells are assumed to be independent and identically exponentially distributed with parameter β.

The following distributional properties (see the appendix and *Smith and Krajewski* [1987]) can be obtained from the model representation of (7):

$$\mu = \lambda\beta^{-1}\pi r^2\gamma\pi a^2 \tag{10}$$

$$\sigma^2 = E\{Y(x)\}(2 + \gamma\pi r^2)/\beta \tag{11}$$

$$p = \exp\{-\lambda\pi a^2[1 - \exp(-\gamma\pi r^2)]\} \tag{12}$$

$$q(s) = \exp\{-\lambda[2\pi a^2 - A(s)][1 - \exp(-\gamma\pi r^2)]\} \tag{13}$$

where

$$A(s) = \pi a^2 - s(a^2 - 0.25s^2)^{1/2} - 2a^2 \arcsin(0.5sa^{-1}) \qquad s \le 2a \tag{14}$$

is the area of overlap of two circles of radius a, whose centers are separated by distance s.

The index of variability for the model is given by

$$I = f(\gamma)$$
$$I = [(2 + \gamma\pi r^2)/\gamma\pi r^2][1 - \exp(-\gamma\pi r^2)] \tag{15}$$

It is straightforward to show that f is monotonically decreasing and that

$$1 < I < 2 \tag{16}$$

The index of variability provides a measure of spatial variability of storm fields that accounts for the spatial intermittency of rainfall. For a "completely random" model of storm fields the index of variability takes the value 2. A completely random model has rain cells of radius r which are distributed throughout the region according to a spatial Poisson process. The rainfall intensity of a rain cell is exponentially distributed. The index of variability takes the value 2, regardless of the spatial rate of occurrence of rain cells or the exponential parameter of rain cell intensity. For the "cluster" model of (8) the index of variability is always less than the value for the completely random model. The index of variability depends only on the spatial rate of occurrence of rain cells within rain bands, γ. If γ is very small the index of variability is close to 2, the value for the completely random model. For large values of γ the index of variability can drop to half the value of the completely random model.

3. PARAMETER ESTIMATION

Parameter estimation for the rainfall model will be based on observations from a single radar and observations from k rain gages with locations $x_1, \cdots, x_k$ relative to the location of the radar. The radar field observations for the ith storm will be denoted $\{Z_i(x)\}$.

It is assumed that radar can accurately distinguish regions receiving rainfall from regions receiving none, that is,

$$Y_i(x) = 0 \qquad \text{if and only if} \quad Z_i(x) = 0 \tag{17}$$

The above assumption is quite accurate given proper processing of the radar data and elimination of ground clutter and anomalous propagation regions. Parameters of the rainfall model are estimated from raingage data and "0-1 mosaics" of the radar field

$$\tilde{Z}_i(x) = 1(Z_i(x) > 0) \tag{18}$$

We will denote by H_n the data set consisting of rain gage observations and 0-1 mosaics of radar fields for n storm fields

$$H_n = \{\tilde{Z}_i(x),\ Y_i(x_j);\ i \le n, \qquad j = 1, \cdots, k;\ x_j \in R\} \tag{19}$$

The rain gage observation for the gage at spatial location x_j and storm i is $Y_i(x_j)$. The total number of rain gages is k. Several assumptions are implicit in the data model of (19). Observations from radar and raingages are time-integrated values covering identical time intervals. Radar observations are continuous in space. It is also assumed that time-integrated rain gage observations are error-free.

Estimation is based on the method of moments (or substitution) principle. The sample moments used in the estimation procedure are the following:

$$\hat{\mu} = n^{-1} \sum_{i=1}^{n} \left[k^{-1} \sum_{j=1}^{k} Y_i(x_j) \right] \tag{20}$$

$$\hat{\sigma}^2 = n^{-1} \sum_{i=1}^{n} \left[k^{-1} \sum_{j=1}^{k} Y_i(x_j)^2 \right] - \hat{\mu}^2 = \hat{v}^2 - \hat{\mu}^2 \tag{21}$$

$$\hat{p} = n^{-1} \sum_{i=1}^{n} \left[|R|^{-1} \int_R [1 - \tilde{Z}_i(x)]\, dx \right] \tag{22}$$

$$\hat{q}(s) = n^{-1} \sum_{i=1}^{n} \left[|R_s|^{-1} \int_{R_s} [1 - \tilde{Z}_i(x)][1 - \tilde{Z}_i(x + y)]\, dx \right] \tag{23}$$

where R_s is the set of points x such that $(x + y) \in R$ and y is a vector of length s. The statistic $\hat{q}(s)$ provides an estimator of the joint probability of zero rainfall at two points separated by distance s. In (21) the estimator is expressed as the difference between $\hat{v}^2$, the estimator of the uncentered second moment, and $\hat{\mu}^2$.

Based on (20)–(22) an estimator for the index of variability is

$$\hat{I} = -\log(\hat{p})(\hat{\sigma}/\hat{\mu})^2 \tag{24}$$

If $\hat{I} \in (1, 2)$ it follows from (15) and (16) that the parameter γ can be estimated as

$$\gamma = f^{-1}(\hat{I}) \tag{25}$$

where the function f is defined in (15). From (12) and (13) we can devise an estimator of the rain band radius by noting that

$$2 - \log q(s)/\log p = A(s)/(\pi a^2) = h(a) \tag{26}$$

where the distance s is specified in (23). It is straightforward to show that

$$0 < h(a) < 1 \tag{27}$$

$$h(a) = 0 \qquad a < s/2 \tag{28}$$

For $a > s/2$ the function $h(a)$ is monotonically increasing. The estimator of $\hat{a}$ is chosen to be

$$\hat{a} = h^{-1}(2 - \log \hat{q}(s)/\log \hat{p}) \tag{29}$$

The estimators for β and λ can now be obtained from (10) and (11) as follows:

$$\hat{\beta}^{-1} = (\hat{\sigma}^2/\hat{\mu})(2 + \hat{\gamma}\pi r^2)^{-1} \tag{30}$$

$$\hat{\lambda} = -\log \hat{p}(\pi \hat{a}^2)(1 - e^{-\hat{\gamma}\pi r^2}) \tag{31}$$

The following proposition establishes asymptotic normality and mean square error consistency of our estimators.

Proposition: If $1 < I < 2$ and $s < 2a$ then there is a covariance matrix Σ such that

$$n^{1/2}(\hat{a}, \hat{\gamma}, \hat{\beta}, \hat{\lambda}) \xrightarrow{D} N((a, \gamma, \beta, \lambda,), \Sigma) \tag{32}$$

where $\xrightarrow{D}$ denotes convergence in distribution and $N((a, \gamma, \beta, \lambda), \Sigma)$ denotes a multivariate normal distribution with mean $(a, \gamma, \beta, \lambda)$ and covariance matrix Σ. We sketch the proof below.

The sample moments $\hat{\mu}$, $\hat{v}^2$, $\hat{p}$, $\hat{q}$ of (20)–(23) can be represented in the form

$$n^{-1} \sum_{i=1}^{n} X_i \tag{33}$$

where the X_i are iid random variables. It follows from the central limit theorem that

$$n^{1/2}(\hat{\mu}, \hat{v}^2, \hat{p}, \hat{q}) \xrightarrow{D} N((\mu, v^2, p, q), \Sigma') \quad (34)$$

for some covariance matrix Σ'. From (25), (29), (30), and (31) we have

$$\hat{a} = g_1(\hat{\mu}, \hat{v}^2, \hat{p}, \hat{q}) \quad (35)$$

$$\hat{\gamma} = g_2(\hat{\mu}, \hat{v}^2, \hat{p}, \hat{q}) \quad (36)$$

$$\hat{\beta} = g_3(\hat{\mu}, \hat{v}^2, \hat{p}, \hat{q}) \quad (37)$$

$$\hat{\lambda} = g_4(\hat{\mu}, \hat{v}^2, \hat{p}, \hat{q}) \quad (38)$$

Let

$$D = [\partial g_i/\partial x_j]_{x=(\mu, v^2, p, q)} \qquad i = 1, \cdots, 4 \qquad j = 1, \cdots, 4 \quad (39)$$

If $g = (g_1, g_2, g_3, g_4)$ has a unique inverse in a neighborhood of (μ, v^2, p, q) then it follows from theorem 3.3.A in the work by *Serfling* [1980] that

$$n^{1/2}(\hat{a}, \hat{\gamma}, \hat{\beta}, \hat{\lambda}) \xrightarrow{D} N((\alpha, \gamma, \beta, \lambda), D\Sigma' D^T) \quad (40)$$

The condition $I \in (1, 2)$ above and the functional form of estimators (35)–(38) ensure that g is invertible in a neighborhood of (μ, v^2, p, q). It follows from the proposition that our estimators are asymptotically unbiased and that variances of the estimators converge to zero. Thus mean square error consistency follows directly from the proposition.

The condition $s < 2a$ in the proposition can be removed if the estimator of (29) for rain band radius is replaced by the following estimator. Choose $\hat{a}$ to minimize the function

$$h(a) = \int_0^{t_0} \left[\frac{\log \hat{q}(s)}{\log \hat{p}} - \frac{2\pi a^2 - A(s)}{\pi a^2} \right]^2 ds \quad (41)$$

where t_0 is constrained by the sampling domain of the radar. Asymptotic normality can be established for the estimator based on (41). The derivation, however, is considerably more complex than for the estimator specified by (29).

The practical utility of the asymptotic properties of parameter estimators is small if similar properties do not hold for realistic sample sizes. This is especially important, since radar records are relatively short compared to rain gage records. Therefore it may be impossible to realize the benefits of multiple sensor rainfall modeling if the parameter estimators do not behave well for small samples. To establish small sample properties of parameter estimators a Monte-Carlo experiment was conducted. This study is described in the following sections.

4. Monte Carlo Experiment

Smith and Krajewski [1987] point out that if the estimated index of variability is outside of the range (1, 2) for a given data set it is impossible to determine whether this is due to inadequate structure of the model or insufficient sample size. A numerical simulation experiment can be helpful to determine the sample size requirements assuming that the latter is the case. In the experiment we generate rainfall fields from our model, then "observe" them by simulating performance of the two sensors: rain gages and radar, and use the sensor-observed data to infer the initial parameters of the underlying rainfall fields.

The selection of parameter values is clearly related to the feasibility issue for the model. The model parameters are unrestricted in the positive space (R_+). The condition imposed on the index of variability, however, constrains the mean, variance, and zero rain probability. Figure 2 presents the feasibility region of our rainfall model in terms of zero rain probability and coefficient of variation. The upper curve corresponds to $I = 1$ and the lower one to $I = 2$. Since both coefficient of variation and zero rain probability are simple statistics with intuitive physical meaning, it is relatively simple to establish the expected range of these parameters for a given geographical location. Zero rain probability is strongly related to the climatology of a region. Similarly, the coefficient of variation characterizes local climate variability. From Figure 2 we can also see that for low probability of rain (high zero rain probability > 0.9) relatively small sampling errors in computations of $\hat{p}$ can put the statistic $\hat{I}$ out of the feasible range.

In order to run our model for any point in the p-I domain we had to determine the corresponding parameters r, a, λ, γ, and β. The parameter γ can be determined from (25) for an assumed radius of a rain cell. In our study we investigated r in the range of 1–2 km^2. Similarly, if one assumes the radius of a rain band one can determine λ from (12). Parameter β can be most conveniently specified from (10), which links the mean rainfall with its average intensity over a cell. Since the concept of a cell is, to a certain degree, a conceptual product it is better to specify the required level of the mean rainfall and adjust parameter β accordingly.

As is clear from the above discussion, a systematic analysis over the range of parameters would lead to a large number of simulation runs. For example, considering 12 pairs of p and I, 2 values for r, 3 for a, and 4 for μ would result in 288 possible combinations. Full analysis of one combination, which includes generation of one thousand storm fields, takes about 6 hours of CPU time on the computer systems available to the authors (Apollo DN3000 and Prime 750). This translates into 72 days of CPU time, a number clearly unmanageable. Therefore only selected runs were performed. The results of these preliminary runs indicated bias in the estimation of the parameters a and I. These biases were fairly independent of the sampling size in the range 25–100.

To further investigate this issue, it was observed that the simulation study could be performed in a "decoupled" manner. There are two types of parameters. One group includes parameters responsible for point characteristics of rainfall, the mean μ, and the variance σ^2. The controlling parameters are β and λ. The second group includes spatial characteristics of the model such as p and is controlled by parameters α, λ, and γ (in the original formulation of *Smith and Krajewski* [1987] these are a_1, a_2, θ, λ and γ). Due to the fact that our model is statistically homogeneous we can compute certain statistics at a point representing a single rain gage. The statistics of interest at a single point must be related to the spatial structure of the model. This condition is met by the average number of cells affecting a single point in space and its associated variance. In addition, the mean $\hat{\mu}$ and the variance $\hat{\sigma}^2$ can be computed. Derivations of the theoretical moments of the number of cells at a point and the corresponding rainfall intensity are given in the Appendix.

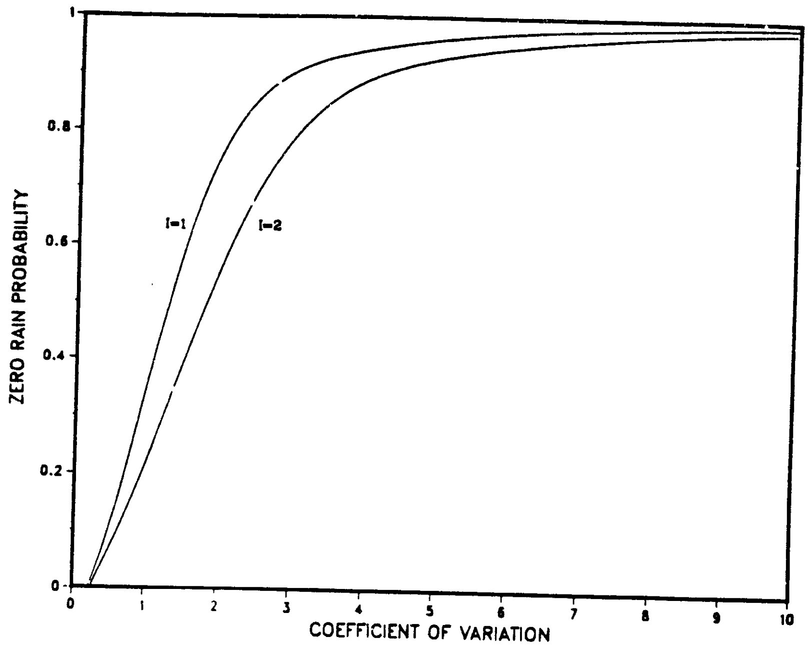

Fig. 2. Feasibility domain of the rainfall model in terms of zero rain probability and coefficient of variation.

The sampling properties of the model at a point were investigated using a simple cell generator. Since the mean number of rainbands at any given point is $\lambda\pi a^2$ and the number of cells at any point within a band is $\gamma\pi r^2$ one can easily generate the cells. An experiment was performed for eight different sample sizes ranging from 25 all the way to 10,000. For each sample size 1000 independent realizations were generated and the mean, the variance, the zero rain probability, the variability index, the average number of cells, and the variance of the average number of cells were computed and compared statistically against the theoretical values. Figure 3 presents typical results of the experiment which was conducted for 72 combinations of the parameters. It is evident that all the statistics, except the variance, display unbiased behavior even for small sample sizes. However, the sampling variability is quite significant. Even for the case with $I = 1.5$ the sampling variability is so high for the sample size of 25, that rejection of the model based on the estimated value of I is quite likely. For those cases where the true value of I lies near the bounds of the feasibility region a sample size of the order of 1000 would be required to ensure high confidence in the adequacy (or inadequacy) of the model. The practical implication of this conclusion depends on the temporal scale of integration of the data. Since the analyzed rainfall fields should be statistically independent, it is estimated that for the daily time scale a record of about 10 years would be required. For the scale of hourly rainfall and a tropical climate probably 3 years of data would be sufficient.

Since the inference of point rainfall characteristics is based on rain gage data, which often are in error, we investigated the influence of the measurement error on the parameter estimates. The error was assumed Gaussian with zero mean and standard deviation of ε percent of the mean. The case of $\varepsilon = 10\%$ was investigated. It was found that the error of 10%, which seems to be a widely accepted error magnitude for rain gage data, has little effect on the mean values of the model parameter estimates. The variances of the estimators are somewhat increased.

With respect to the spatial parameters of the model, their estimation is based on the function $h(\ )$ given by (41). Computation of the logarithmic function $\log \hat{q}(s)$ is the most time consuming component of the parameter estimation algorithm. Since $\hat{q}(s)$ depends on several parameters such as λ, γ, r, and a, a comprehensive evaluation of the sampling properties of the spatial parameters was not feasible. However, a simplified scenario was analyzed in detail. Due to the fact that our main concern is the "ability" of the two-point zero rain probability function to reveal the size of the rain bands we performed a one-dimensional experiment addressing this question.

Consider an arbitrary cross section of the radar domain passing through the center. If we wanted to estimate the band radius based on the two-point zero rain probability function computed along such a line we would face a problem resulting from the fact that, in general, our line does not pass through the centers of all the bands that it is crossing. However, the one-dimensional equivalent of the model investigated in this work requires that the line segments corresponding to the bands be of the same length. Similarly, the line segments corresponding to rain cells are also of fixed length. The line analogy of the spatial model is presented in Figure 4. To investigate the ability of $\hat{q}$ to estimate the rainbands radius we performed a Monte-Carlo simulation over a range of sampling conditions. The length of the line was taken as 100. The intensity parameter λ_l controls the number of "bands" within the line. Parameter γ_l controls the number of "cells" within a band. "Cell" radius is r_l and band radius is a_l. It can be easily shown that

$$h_l(a_l) = \int_0^{t_l} \left[\frac{\log \hat{q}_l(s)}{\log \hat{p}_l} - \frac{4a_l - A_l(s)}{2a_l} \right]^2 ds \qquad (42)$$

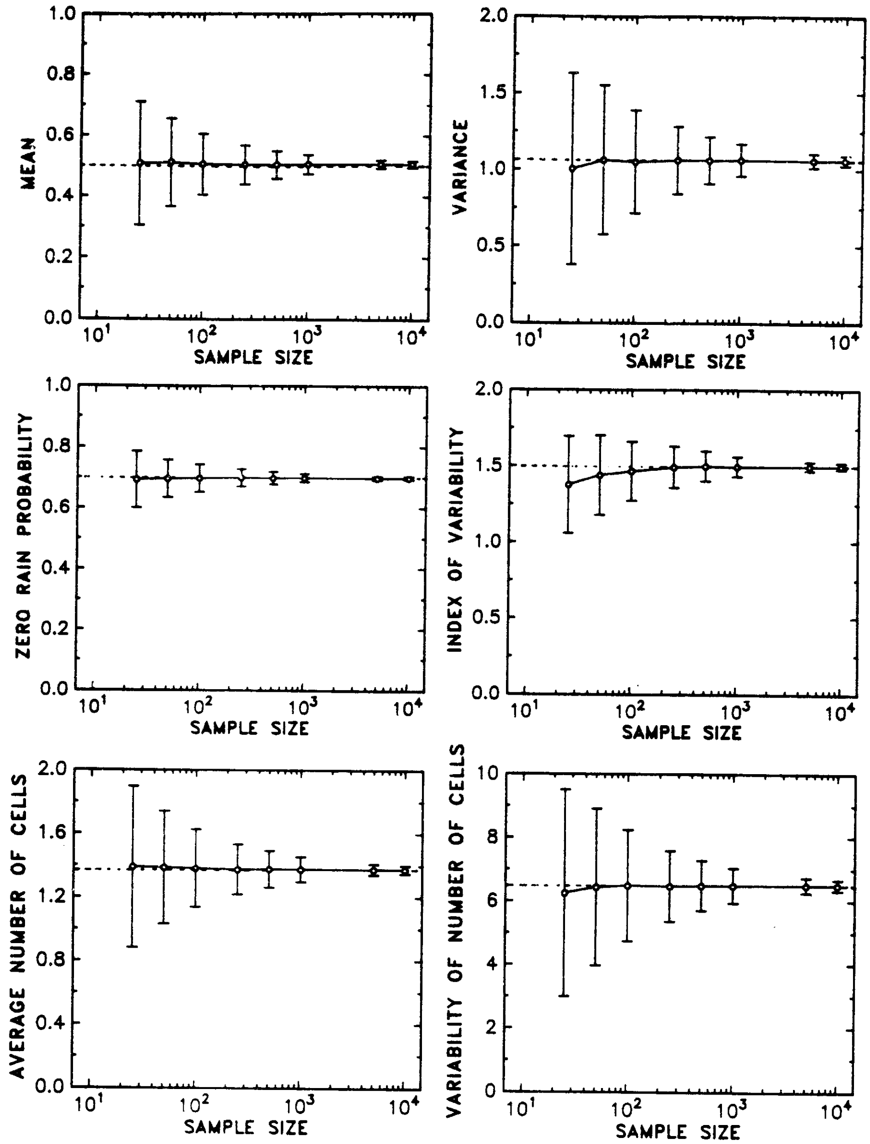

Fig. 3. Results of simulations of a point cell generator. The true values of the statistics are indicated to be the dashed lines. The results are for $\lambda = 0.005$, $\gamma = 1.2$, $\beta = 2.7$, $a = 5.0$, and $r = 1.0$. The horizontal bars indicate one standard deviation range.

where

$$A_l(s) = 2a_l - s \qquad s \leq 2a_l$$
$$A_l(s) = 0 \qquad \text{otherwise} \tag{43}$$

and s is, as before, the distance between the centers of two bands.

Since the magnitude of rainfall is of no concern in estimating the band radius, the line process can be conveniently generated as a 0-1 process. The two-point zero rain probability function can be rapidly computed for such processes.

The first question that we would like to address concerns the limit of λ for which we can expect to be able to estimate the spatial features of rainfall. It is intuitively clear that distinct bands (relatively low λ and high γ) should be the easiest to identify. For values larger than a critical value λ_{cr} the bands will all merge together making estimation of their size impossible. This is illustrated in Figure 5. Similarly, even if λ is under the critical value but γ is low, the bands may appear smaller than they really are. Figure 6 presents the results of our simulations illustrating such a situation. It is evident that for $\lambda > \lambda_{cr}$ the estimation results are becoming erratic and unreliable. As far as the effect of γ is concerned, we can see that severe underestimation occurs until γ reaches about 2. It is also clear that γ can reach a saturation level γ_s. For $\gamma > \gamma_s$ there is no effect on estimation of a_l. But, of course, λ and γ jointly determine the properties of the estimators of interest. We have also investigated the effect of sample size. Over a wide range of parameters the behavior of the estimator $\hat{a}_l$ is remarkably stable in the mean. The variance also is quite small even for samples of the order of 100 (Figure 7).

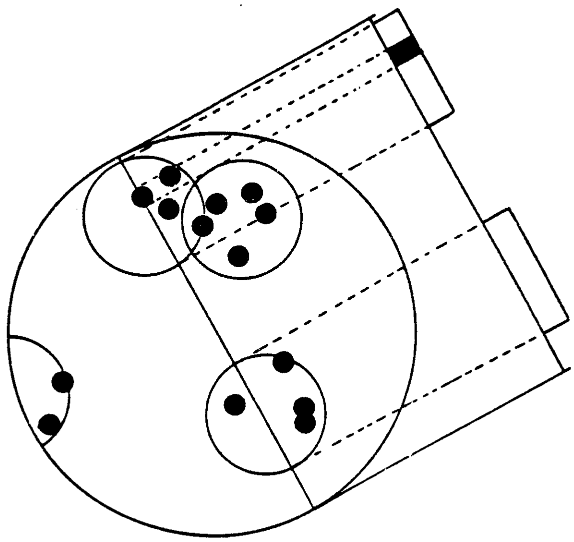

Fig. 4. Schematic representation of line analogy.

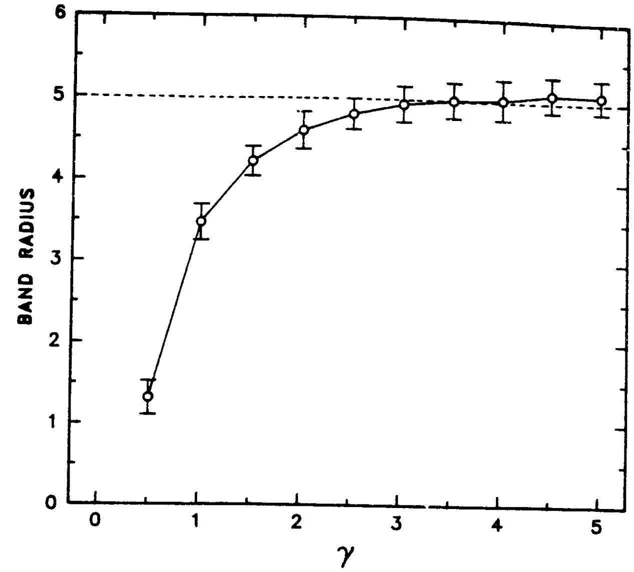

Fig. 6. Results of simulations of a line generator. Shown is influence of cell intensity γ (with $\lambda = 0.1$) on band radius estimation. The true band radius was fixed at 5.0 and cell radius was 1.0. The horizontal bars indicate one standard deviation range.

5. CONCLUSIONS

We have presented investigations of the parameter estimation procedure for a statistical space-time model of rainfall based on raingage and radar observations. In the model, the radar data are used as an indicator of rainy versus dry areas. Therefore only the spatial characteristics of the modeled rainfall fields are inferred from radar data. The parameter estimators for the model possess attractive characteristics; in section 3 we have shown that the estimators are asymptotically normal. the central question of this paper was to establish data requirements for parameter estimation. To accomplish this goal we need to asses the small sample properties of parameter estimates. The approach undertaken was that of a Monte-Carlo study. We generated data records, for both radar and rain gages, from our model and then tried to estimate back the parameters of the model. Due to the number of parameters of the model, a comprehensive evaluation of the behavior of the model parameters throughout the whole parameter space was not feasible and we had to decompose the model into blocks which could be investigated independently. There were two major blocks: first, we evaluated the ability of the model to infer point rainfall magnitude from rain gage data; second, we evaluated the model's ability to estimate the spatial features of the model. The first issue was resolved positively by means of our experiment. For small samples, say, less than 100 independent events, bias is a clear problem. For samples on the order of 500–1000 the estimators are effectively unbiased and characterized by small variances. The second issue was answered less than satisfactorily. The parameter space which is useful in terms of the model's ability to estimate the spatial characteristics of rainfall patterns is restricted. Parameters being outside of this space result in biased results. On a positive note, the sample size required to estimate the spatial characteristics is smaller than that required to estimate the magnitude of point rainfall. Now, a very important reminder is due. The above conclusions are based on the assumption that the model is correct in the sense that the rainfall fields are organized in exactly the way which was proposed in the model. Of course, within our Monte Carlo experiment this assumption was met, however, in reality parametric model assumptions are seldom perfectly satisfied.

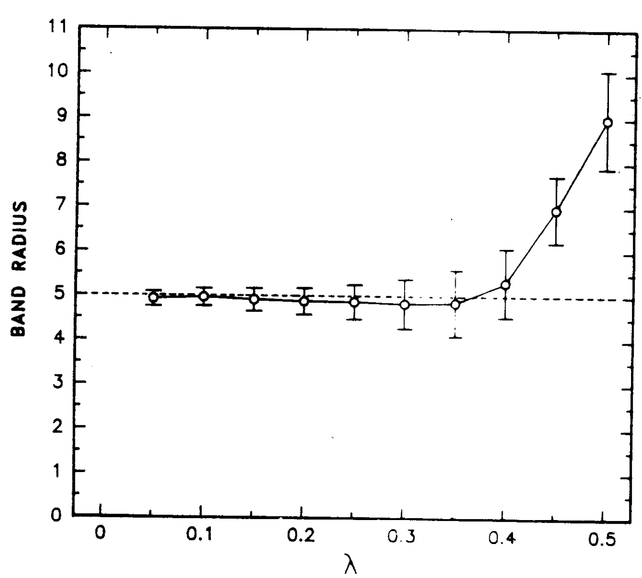

Fig. 5. Results of simulations of a line generator. Shown is influence of band intensity λ (with $\gamma = 3.0$) on band radius estimation. The true band radius was fixed at 5.0 and cell radius was 1.0. The horizontal bars indicate one standard deviation range.

From a practical point of view, if the variability index says that the model cannot be used, it is really immaterial whether the reason is small sample size or inadequate model structure. In both cases one has to look for alternative approaches. However, the results of our simulations tend to indicate that with about 500 statistically independent rainfall fields one should be able to avoid most of the sampling

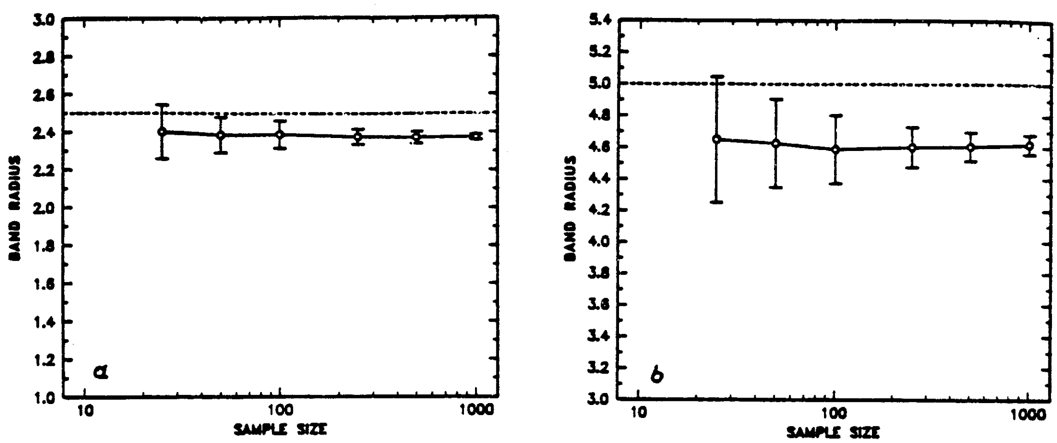

Fig. 7. Results of simulations of a line generator. The model parameter values are (*a*) $\lambda = 0.05$, $\gamma = 2.0$, $a = 2.5$, and $r = 1.0$ and (*b*) $\lambda = 0.10$, $\gamma = 2.0$, $a = 5.0$, and $r = 1.0$. The horizontal bars indicate one standard deviation range.

problems. In order to translate this number into an actual sample size one needs to consider time correlation length and climatology of a particular location. Whether the anisotropy assumption violates these findings is another issue not addressed here.

APPENDIX

In this appendix we present derivations of the moment equations for the model presented in section 2. The representation

$$Y(x) = \sum_{i=1}^{N(x)} U_i \tag{A1}$$

is used to obtain (10)–(11):

$$\begin{aligned} E\{Y(x)\} &= E\left\{E\left\{\sum_{i=1}^{N(x)} U_i | N(x)\right\}\right\} \quad \text{(Wald's lemma)} \\ &= E\{N(x)E\{U_i\}\} \\ &= \beta^{-1}E\{N(x)\} \end{aligned} \tag{A2}$$

because U_i has an exponential distribution with parameter β. Similarly,

$$\begin{aligned} \text{Var}\,\{Y(x)\} &= E\{N(x)\}\,\text{Var}\,\{U_i\} + \text{Var}\,\{N(x)\}E\{U_i\}^2 \\ &= \beta^{-2}[E\{N(x)\} + \text{Var}\,\{N(x)\}] \end{aligned} \tag{A3}$$

The number of cells $N(x)$ can be represented as follows. If M rain bands cover the point x, the number of cells $N(x)$ has a Poisson distribution with mean $M\gamma\pi r^2$ (γ is spatial rate of occurrence of cells; r is cell radius). The number of rain bands M covering x has a Poisson distribution with mean $\lambda\pi a^2$ (λ is spatial rate of occurrence of bands; a is radius of a band. Therefore

$$\begin{aligned} E\{N(x)\} &= E\{E\{N(x)|M\}\} \\ &= E\{M\gamma\pi r^2\} \\ &= \lambda\pi a^2\gamma\pi r^2 \\ &= \bar{\lambda}\bar{\gamma} \end{aligned} \tag{A4}$$

with $\bar{\lambda} = \lambda\pi a^2$ and $\bar{\gamma} = \gamma\pi r^2$.

Also,

$$\begin{aligned} E\{N^2(x)\} &= E\{E\{N^2(x)|M\}\} \\ &= E\{M\bar{\gamma} + (M\bar{\gamma})^2\} \\ &= \bar{\lambda}\bar{\gamma} + \bar{\gamma}^2(\bar{\lambda} + \bar{\lambda}^2) \end{aligned} \tag{A5}$$

Thus the variance

$$\begin{aligned} \text{Var}\,\{N(x)\} &= \bar{\lambda}(\bar{\gamma} + \bar{\gamma}^2) \\ &= \lambda\pi a^2[\gamma\pi r^2 + (\gamma\pi r^2)^2] \end{aligned} \tag{A6}$$

Now

$$\begin{aligned} E\{Y(x)\} &= \beta^{-1}E\{N(x)\} \\ &= \beta^{-1}\bar{\lambda}\bar{\gamma} \\ &= \beta^{-1}\lambda\pi a^2\gamma\pi r^2 \end{aligned} \tag{A7}$$

$$\begin{aligned} \text{Var}\,\{Y(x)\} &= \beta^{-2}[\bar{\lambda}\bar{\gamma} + \bar{\lambda}(\bar{\gamma} + \bar{\gamma}^2)] \\ &= (\beta^{-1}\bar{\lambda}\bar{\gamma})\beta^{-1}(2 + \bar{\gamma}) \\ &= E\{Y\}\beta^{-1}(2 + \gamma\pi r^2) \end{aligned} \tag{A8}$$

Derivations of (12)–(13) are straightforward based on the assumption that the number of cells at a point x, $N(x)$, is Poisson distributed.

Acknowledgments. The research leading to this paper was supported by the Iowa Institute of Hydraulic Research of The University of Iowa and the Office of Hydrology of the National Weather Service. This support is gratefully acknowledged. The authors would like to thank Mike Kundert and Venkataraman Lakshmi who helped with the graphics.

REFERENCES

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

Battan, L. J., *Radar Observation of the Atmosphere*, 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.*, *92*(D8), 9631–9644, 1987.

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

Doviak, R. J., and D. S. Zrnic, *Doppler Radar and Weather Observations*, Academic, San Diego, Calif., 1984.

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

Hudlow, M. D., and V. Patterson, GATE radar rainfall atlas, EDIS, NOAA special report, pp. 155, Dept. of Commerce, Washington, D. C., 1979.

Kavvas, M. L., and K. R. Herd, A radar-based stochastic model for short-time-increment rainfall, *Water Resour. Res.*, *21*(9), 1437–1455, 1985.

Rodriguez-Iturbe, I., V. K. Gupta, and E. Waymire, Scale considerations in the modeling of temporal rainfall, *Water Resour. Res.*, *20*(11), 1611–1619, 1984.

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 and Sons, New York, 1980.

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

Smith, J. A., and W. F. Krajewski, Statistical modeling of space-time rainfall using radar and rain gage observations, *Water Resour. Res.*, *23*(10), 1893–1900, 1987.

Valdes, J. B., I. Rodriguez-Iturbe, and V. K. Gupta, Approximations of temporal rainfall from a multidimensional model, *Water Resour. Res.*, *21*(8), 1259–1270, 1985.

W. F. Krajewski, Iowa Institute of Hydraulic Research, University of Iowa, Iowa City, IA 52242.

J. A. Smith, Hydrologic Research Laboratory, National Weather Service, 8060 13th Street Silver Spring, MD 20910.

(Received January 17, 1989;
revised May 30, 1989;
accepted June 6, 1989.)