WATER RESOURCES RESEARCH, VOL. 21, NO. 8, PAGES 1251–1257, AUGUST 1985

Parameter Estimation for a Model of Space-Time Rainfall

JAMES A. SMITH

Interstate Commission on the Potomac River Basin, Rockville, Maryland

ALAN F. KARR

Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, Maryland

In this paper, parameter estimation procedures, based on data from a network of rainfall gages, are developed for a class of space-time rainfall models. The models, which are designed to represent the spatial distribution of daily rainfall, have three components, one that governs the temporal occurrence of storms, a second that distributes rain cells spatially for a given storm, and a third that determines the rainfall pattern within a rain cell. Maximum likelihood and method of moments procedures are developed. We illustrate that limitations on model structure are imposed by restricting data sources to rain gage networks. The estimation procedures are applied to a 240-mi^2 (621 km^2) catchment in the Potomac River basin.

1. INTRODUCTION

Development of probabilistic models of space-time rainfall has had a lengthy but sporadic history, beginning with *LeCam* [1961]. Research has been motivated mainly by applications, most notably management of water resource systems [*LeCam*, 1961], design of hydrologic monitoring networks [*Rodriguez-Iturbe and Mejia*, 1974; *Bras and Rodriguez-Iturbe*, 1976], design of flood control structures [*Corotis*, 1976], and development of air quality standards [*Rodhe and Grandell*, 1981]. In recent years, increasing emphasis has been placed on the physical basis of rainfall representations [*Eagleson*, 1984; *Gupta and Waymire*, 1979; *Smith and Karr*, 1983; *Waymire et al.*, 1984]. In an assessment of precipitation research prepared by the Precipitation Committee of the American Geophysical Union [*AGU*, 1984] it is noted that

> although physical/dynamical considerations are expected to play an important role in advances in stochastic modeling of precipitation, the problems of parameter estimation and statistical inference are not expected to be solved by appealing only to precipitation physics. New statistical techniques remain to be developed, particularly for stochastic descriptions of space-time rainfall.

In this paper we examine the role of statistical parameter estimation in modeling space-time rainfall. Specifically, we develop and illustrate estimation procedures, based on data from a network of rainfall gages, for a class of space-time rainfall models.

The models we examine are designed to represent the spatial distribution of daily summer season rainfall in a humid region. The models have three basic components, one that governs the temporal occurrence of storms, a second that distributes "rain cells" spatially for a given storm, and a third that determines the rainfall pattern within each rain cell. Construction of the model by successive randomization of model components (using rain cells as basic building blocks) is analogous to the approach taken by *LeCam* [1961] [see also *Waymire et al.*, 1984]. Estimation procedures are developed for a model with five parameters.

We restrict consideration to precipitation data from operational rain gage networks because of availability, quality, and record length of the data. Additional sources of precipitation data include dense experimental networks of rain gages, radar, and satellite imagery. Applicability of models and estimation procedures that rely on these data sources is limited not only by data availability but also by lack of methods for dealing simultaneously with qualitatively different kinds of data. We illustrate in this paper, however, that serious limitations on model structure are imposed by restricting data sources to operational rain gage networks. These considerations suggest that estimation procedures for space-time rainfall models ideally should utilize a combination of "physically determined" parameters and parameters estimated from data (such a procedure is illustrated in section 5).

It should be emphasized that the models under consideration are intended for application to management and design problems, not problems of short-term precipitation prediction. The latter class of problems relies on operational meteorological observations [see, for example, *Georgakakos and Bras*, 1984].

Contents of the sections are as follows. Section 2 contains a description of the rainfall model, including distributional results required for estimation procedures. Maximum likelihood estimators are derived in section 3; method of moments estimation is the subject of section 4. Section 5 contains data analysis results (using daily precipitation data from the Potomac River basin) that illustrate suitability of model assumptions; implementation of the method of moments estimation procedure is also illustrated. Results and conclusions are summarized in section 6.

2. MODEL DEVELOPMENT

Prior to giving a detailed mathematical description of the rainfall model, we give a simplified description of its basic components. Temporal evolution of the model is governed by a "stochastic climatological process" $\{X(n)\}$, where $X(n)$ represents the climatological state on day n (days can be replaced by shorter but homogeneous time steps). The process alternates between two states, 0 and 1, the former representing anticyclonic conditions in which precipitation cannot occur,

Paper number 5W0290.
0043-1397/85/005W-0290$05.00

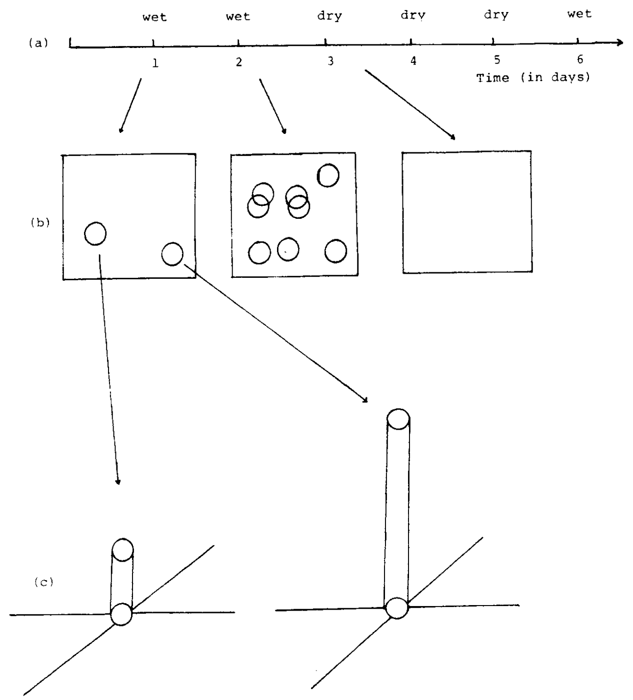

Fig. 1. Schematic of rainfall model components: (*a*) sequence of wet-dry days, (*b*) spatial distribution of rain cells for the first three days, and (*c*) rainfall depths for the two cells on day 1.

and the latter representing conditions in which air-mass thunderstorms predominate.

For days on which rainfall can occur a random "spatial intensity process" determines the spatial rate of occurrence of rain cells. Rain cells are taken to be circular and of fixed radius. Given the spatial rate v of rain cells, the locations of rain cell centers constitute a spatial Poisson process with rate v.

Given the locations of rain cells the final component determines the "rainfall distribution" within individual rain cells. We examine two "canonical" rainfall distribution patterns, cells with uniform rainfall intensity, and cells with rainfall intensity decreasing linearly with distance from the center of the cell. Maximum rainfall intensity occurs at the center of a cell and varies randomly from cell to cell. Figure 1 illustrates the three model components.

Independence assumptions are ubiquitous in the rainfall model. Rather than state explicitly all of the independence assumptions we ask the reader to assume independence unless explicitly stated otherwise. Computations, which typically involve repeated application of Bayes' theorem, are generally omitted; details are available from the authors.

The rainfall model we develop is restricted to a subset R of the plane. The primary quantities of interest are the total rainfall on day n in a subset A of R, $Z_n(A)$ (with units L^3), and the rainfall intensity on day n at point x, $Y_n(x)$ (with units L), which are related by

$$Z_n(A) = \int_A Y_n(x)\, dx \qquad (1)$$

The "climatological process" $\{X(n),\ n \geq 0\}$ is a Markov chain with state space $\{0, 1\}$. Two parameters are introduced into the rainfall model, the probability of transition from a dry state to a dry state

$$q(0) = P\{X(n) = 0 \mid X(n-1) = 0\} \qquad (2)$$

and the probability of transition from a wet state to a wet state,

$$q(1) = P\{X(n) = 1 \mid X(n-1) = 1\} \qquad (3)$$

These determine the limit probabilities

$$p(1) = \lim_{n \to \infty} P\{X(n) = 1\} \qquad (4)$$

$$p(0) = \lim_{n \to \infty} P\{X(n) = 0\} = 1 - p(1) \qquad (5)$$

in particular [see *Cinlar*, 1975],

$$p(1) = (1 - q(0))/(2 - q(0) - q(1)) \qquad (6)$$

Note also that if one of the transition probabilities is known and the limit probabilities are known then the remaining transition probability can be computed from (6). In the method of moments procedure developed in section 4, $p(1)$ and $q(1)$ are estimated directly; $q(0)$ is obtained from (6).

For "rainy days," i.e., days for which $X(n) = 1$, the spatial distribution of rainfall is represented by the superposition of a random number of rain cells of fixed radius r. The "spatial rate" of occurrence of rain cells (the mean number of rain cells per unit area) varies randomly over time (day to day) to reflect the variation in intensity of convective activity; the spatial rate is represented by a nonnegative stochastic process $\{V(n),\ n \geq 0\}$ satisfying

$$P\{V(n) = 0 \mid X(n) = 0\} = 1 \qquad (7)$$

$$P\{V(n) \leq y \mid X(n) = 1\} = F(y) \qquad (8)$$

where F is a distribution function on $[0, \infty)$. The first condition states that on "dry" days the spatial rate is 0. Equation (8) states that the distribution of the rate on "wet" days is given by F. For the estimation procedures developed in the following section we assume that spatial rates are exponentially distributed:

$$F(y) = 1 - \exp(-ay) \qquad (9)$$

Given the spatial rate $V(n)$ on a wet day, the cell center locations form a Poisson process on R with intensity $V(n)$, i.e.,

$$P\{N_n(A) = k \mid V(n)\} = \exp(-V(n)|A|)(V(n)|A|)^k/k! \qquad (10)$$

where $N_n(A)$ is the number of cell centers located in a subset A of R, and $|A|$ is the area of A. Moreover, numbers of cells in disjoint spatial regions are conditionally independent given $V(n)$. For computational purposes, the crucial property of the cell location process is that it is a Cox process (or doubly stochastic Poisson process) on R. Spatial Cox processes are described in the work by *Matern* [1960], while *Smith and Karr* [1983] illustrate the utility of Cox processes for modeling temporal rainfall occurrences. The cell location process introduces two more parameters to the rainfall model, the cell radius r, and the mean spatial rate a^{-1}.

We make two basic assumptions on the rainfall distribution pattern of a rain cell. First, we assume that isohyets of rainfall intensity are circular with maximum intensity at the center of the cell. The second assumption is that total rainfall from a cell is the product of the rainfall intensity at the center of the cell times the volume of the "canonical rainfall distribution," which is specified by a descreasing function h on $[0, \infty)$ satisfying (recall that r is the cell radius)

$$0 \le h(s) \le 1$$
$$h(s) = 0 \qquad s > r$$
$$h(0) = 1$$

The function h evaluated at s specifies the rainfall intensity at distance s from the center of the cell for a storm with unit maximum intensity. The total rainfall from a rain cell with unit maximum intensity is thus given by

$$m = 2\pi \int_0^r sh(s)\, ds \tag{11}$$

We illustrate with two special cases.

2.1. *Example 1*

Uniform rainfall intensity is represented by

$$h(s) = 1 \qquad s \le r \tag{12}$$

Total rainfall from a cell with uniform rainfall intensity and unit maximum intensity is πr^2.

2.2. *Example 2*

Linearly decreasing rainfall isohyets are represented by

$$h(s) = 1 - s/r \qquad s \le r \tag{13}$$

Total rainfall from a cell with linearly decreasing isohyets and unit maximum intensity is $(1/3)\pi r^2$.

Hereafter we assume that the canonical rainfall distribution is the uniform rainfall intensity of example 1.

The maximum intensity for cell k on day n, $I(n, k)$, is a random variable with distribution G, assumed below to be exponential, i.e.,

$$P\{I(n, k) \le y\} = 1 - \exp(-by) \tag{14}$$

With this assumption the rainfall distribution pattern adds one parameter to the rainfall model, the mean value b^{-1} of maximum rainfall intensity for a cell.

The total rainfall in a subset A of R contributed by a cell (with unit maximum intensity) located at the origin is given by

$$m(A) = \int_A sh(s)\, d\theta\, ds \tag{15}$$

The total rainfall in A contributed by the kth cell on day n is hence $I(n, k)m(A - L(n, k))$, where $L(n, k)$ is the location of the kth cell on day n, and $A - L(n, k)$ is the translation of the set A by $L(n, k)$, i.e.,

$$A - L(n, k) = \{x \in R : x + L(n, k) \in A\}$$

The total rainfall in A on day n is the sum of contributions from all cells on day n

$$Z_n(A) = \sum_{k=1}^{N_n(R)} I(n, k)m(A - L(n, k)) \tag{16}$$

The rainfall intensity at a point x in R is given by

$$Y_n(x) = \sum_{k=1}^{N_n(R)} I(n, k)h(|x - L(n, k)|) \tag{17}$$

where $|x - L(n, k)|$ is the distance from x to the location of cell k on day n.

The distributional result that forms the basis for the estimation procedures developed below is the following expression for the joint Laplace transform of total rainfall in k disjoint sets. The proof is sketched in the appendix.

2.3. *Proposition*

For disjoint sets $A_1, \cdots, A_k \subset R$ and nonnegative $s_1, \cdots s_k$,

$$E\left[\exp\left(-\sum_{i=1}^k s_i Z_n(A_i)\right)\right] = [p(1)]\left[1 + a^{-1}\int_R \left[1 - b\left\{b + \sum_{i=1}^k s_i m(A_i - x)\right\}^{-1}\right] dx\right]^{-1} + (1 - p(1)) \tag{18}$$

Of more direct use in obtaining estimators is the following representation of rainfall intensity at k points that are sufficiently far apart.

Corollary

For $x_1, \cdots, x_k \in R$ such that the minimum distance between any two points is greater than $2r$,

$$E\left[\exp\left(-\sum_{i=1}^k s_i Y_n(x_i)\right)\right] = [p(1)]\left[1 + a^{-1}\pi r^2\left[k - \sum_{i=1}^k b(b + s_i)^{-1}\right]\right]^{-1} + (1 - p(1)) \tag{19}$$

The following distributional results can be obtained from (18) and (19):

$$E[Y_n(x)] = p(1)b^{-1}a^{-1}\pi r^2 \tag{20}$$

$$\mathrm{Var}\,(Y_n(x)) = p(1)(2 - p(1))b^{-2}a^{-2}\pi^2 r^4 + 2p(1)b^{-2}a^{-1}\pi r^2 \tag{21}$$

$$\mathrm{Cor}\,(Y_n(x), Y_n(y)) = \frac{1 + \dfrac{H(|x - y|)}{(2 - p(1))a^{-1}\pi^2 r^4}}{1 + \dfrac{2}{(2 - p(1))a^{-1}\pi r^2}} \tag{22}$$

where $|x\text{-}y|$ is the distance from x to y, and

$$\begin{aligned} H(x) &= 4r^2 \cos^{-1}(x/2r) - x(r^2 - (x/2)^2)^{1/2} \qquad x \le 2r \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{23}$$

$$P\{Y_n(x) = 0 \mid X(n) = 1\} = 1/(1 + a^{-1}\pi r^2) \tag{24}$$

$$P\{Y_n(x) \le y \mid X(n) = 1\} = (1 - B)A \exp\{-Ay\} + B1(y = 0) \tag{25}$$

where

$$A = b/(1 + a^{-1}\pi r^2)$$
$$B = 1/(1 + a^{-1}\pi r^2)$$

and

$$\begin{aligned} 1(y = 0) &= 1 \qquad y = 0 \\ &= 0 \qquad \text{otherwise} \end{aligned}$$

Equation (19) has important implications for parameter estimation (note that equation (19) applies when the rain gage density is "sparse" in the sense that the minimum distance between gages is larger than the cell diameter). Since the parameters a and r appear only through the product $a^{-1}r^2$, it will be impossible to estimate a and r individually. In statistical terms the five-parameter model is not identifiable.

It follows from (25) that the distribution of rainfall intensity at a gage on a wet day is zero with positive probability. On "wet" days for which rainfall is recorded at the gage, the distribution of rainfall intensity is exponential.

3. MAXIMUM LIKELIHOOD ESTIMATION

In this section we develop a maximum likelihood procedure for estimating parameters of the model described in the previous section. The estimation procedures (for this section and the following section on method of moments estimation) are based on data from a network of k rain gages. We denote the observed rainfall at gage j on day n by $Y_n(j)$. The vector $(Y_n(1), \cdots, Y_n(k))$ of observations on day n will be denoted $\bar{Y}_n$.

Geometry of the rain gage network enters into the estimation procedures only through the minimum distance between gages. In particular, we assume that the rain cell diameter $2r$ is smaller than the minimum rain gage distance d; this assumption implies that the network is not dense enough for two gages to observe the same rain cell. Under this assumption it is possible to estimate only four parameters: $q(0)$, $q(1)$, b, and $a^{-1}r^2$, as was discussed in the previous section. To obtain estimates of either a or r it is necessary that one of the parameters be estimated by additional data sources. It is much more likely for exogenous estimates of the cell radius r to be available than estimates of the mean number of cells a^{-1}. Consequently, we assume hereafter that r is "known."

The sample function density $f(\bar{y}_1, \cdots, \bar{y}_n)$ for observations at k gages over n days is the joint probability density function of the n daily observation vectors $\bar{y}_1, \cdots, \bar{y}_n$:

$$P\{\bar{Y}_1 \in A_1, \cdots, \bar{Y}_n \in A_n\} = \int_{A_1} \cdots \int_{A_n} f(\bar{y}_1, \cdots, \bar{y}_n)\, d\bar{y}_1 \cdots d\bar{y}_n \tag{26}$$

where $A_i \in R_+{}^k$. Maximum likelihood estimation is based on the likelihood function

$$1(\theta \mid \bar{y}_1, \cdots, \bar{y}_n) = f(\bar{y}_1, \cdots, \bar{y}_n) \tag{27}$$

where $\theta = (q(0), q(1), a, b)$ is the vector of unknown parameters. The notation emphasizes that in estimation problems the data are treated as fixed and the parameters as variable.

We show below that computation of the likelihood function can be broken into two parts, one involving the temporal evolution of the model, and the other involving the spatial distribution of rainfall on a wet day. We first note that the likelihood function can be written in "iterated form" as

$$1(\theta \mid \bar{y}_1, \cdots, \bar{y}_n) = \prod_{k=2}^{n} g(\bar{y}_k \mid \bar{y}_{k-1}, \cdots, \bar{y}_1) g(\bar{y}_1) \tag{28}$$

where $g(\bar{y}_k \mid \bar{y}_{k-1}, \cdots, \bar{y}_1)$ is the conditional density of $\bar{Y}_k$ given $\bar{Y}_{k-1}, \cdots, \bar{Y}_1$.

The computational split into spatial and temporal components is used for computation of the conditional density g:

$$g(\bar{y}_n \mid \bar{y}_{n-1}, \cdots, \bar{y}_1) = h(\bar{y}_n) P\{X(n) = 1 \mid \bar{y}_{n-1}, \cdots, \bar{y}_1\} + 1(\bar{y}_{n=0}) P\{X = (n) = 0 \mid \bar{y}_{n-1}, \cdots, y_1\} \tag{29}$$

where h is the conditional density of $\bar{Y}_n$ given $X(n) = 1$ and

$$\begin{aligned} 1(\bar{y}_n = 0) &= 1 \quad \text{if } y_n(i) = 0 \qquad i = 1, \cdots, k \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{30}$$

is the indicator function for the event that rainfall is not recorded at any gage on day n. The conditional density h thus describes the distribution of rainfall at the k gages on wet days, while $P\{X(n) = i \mid \bar{y}_{n-1}, \cdots, \bar{y}_1\}$ (for i equal to 0 or 1) summarizes information on temporal evolution of the climatological process contained in the preceding rain gage observations. Below we show how the "spatial component" and "temporal component" of (29) can be computed. We begin with $P\{X(n) = 1 \mid \bar{y}_{n-1}, \cdots, \bar{y}_1\}$.

It follows from the Markov property of $\{X(n)\}$ that

$$P\{X(n) = 1 \mid \bar{Y}_{n-1}, \cdots, \bar{Y}_1\} = P\{X(n) = 1 \mid T(n)\} \tag{31}$$

where

$$T(n) = n - \max\{i : \bar{Y}_i \neq 0, i \leq n-1\} \tag{32}$$

is the number of days since rainfall has been recorded at any gage. Consequently, it suffices to compute

$$s(j) = P\{X(n) = 1 \mid T(n) = j\} \tag{33}$$

for which we obtain the following recursive

$$s(1) = q(1) \tag{34}$$

for $j > 1$,

$$s(j) = [1 - q(0)] + [q(0) + q(1) - 1] \cdot \frac{s(j-1)}{1 + ka^{-1}\pi r^2[1 - s(j-1)]} \tag{35}$$

The conditional density h can be described by its Laplace transform

$$\begin{aligned} K_n(s_1, \cdots, s_k) &= E[\exp\{-s_1 Y_n(1) - \cdots - s_k Y_n(k)\} \mid X(n) = 1] \\ &= \int_0^\infty \cdots \int_0^\infty \exp\{-s_1 y_1 - \cdots - s_k y_k\} h(\bar{y})\, dy_1 \cdots dy_k \end{aligned} \tag{36}$$

In particular, from (19) we obtain

$$K_n(s_1, \cdots, s_k) = \left[1 + a^{-1}\pi r^2\left(k - \sum_{i=1}^{k} b(b + s_i)^{-1}\right)\right]^{-1} \tag{37}$$

While direct evaluation of h by inversion of (37) appears difficult, (37) can be used either for numerical evaluation of h or as the basis of approximations.

4. METHOD OF MOMENTS ESTIMATION

The framework for parameter estimation in this section is the same as in the previous section. Specifically, we assume that precipitation data $(\bar{Y}_1, \cdots, \bar{Y}_n)$ are available for n days from k gages. The method of moments procedure shares the characteristic of the maximum likelihood procedure that estimation separates into spatial and temporal components. In this case separation is achieved through an iterative computation procedure with two steps to each iteration: (1) the spatial parameters a and b are estimated from the sample mean and variance with the temporal parameters set equal to their values at the previous iteration, and (2) the temporal parameters are estimated from the sequence of wet and dry days with the spatial parameters set equal to the values obtained in step 1. The procedure is terminated when the relative change in each parameter is less than a prescribed value.

Step 1 is based on (20) and (21), which present the mean and variance of rainfall intensity at a site. The following estimators are obtained for a and b:

$$\hat{b} = 2\hat{\mu}/(\hat{\sigma}^2 - [(2 - \hat{p}(1))/\hat{p}(1)]\hat{\mu}^2) \quad (38)$$

$$\hat{a} = \hat{p}(1)\pi r^2/(\hat{\mu}\hat{b}) \quad (39)$$

where $\hat{\mu}$ and $\hat{\sigma}^2$ are the sample mean and variance.

The estimator of $p(1)$, the stationary probability of being in a wet state on a given day, is the sum of two terms: (1) the fraction of days for which rainfall was recorded at one or more gages and (2) the fraction of days for which no rainfall was recorded at any gage multiplied by the conditional probability of being in the wet state given that no rainfall was recorded. This estimator can be represented as

$$\hat{p}(1) = (1/n) \sum_{i=1}^{n} 1(\bar{y}_i \neq 0) + P\{X(i) = 1 \mid \bar{Y} = 0\}(1/n) \sum_{i=1}^{n} 1(\bar{y}_i = 0) \quad (40)$$

Dependence on spatial parameters is through the conditional probability of being in the wet state given that no rainfall was recorded, for which we have the expression

$$P\{X(i) = 1 \mid \bar{Y}_i = 0\} = (1 + \{(1 - p(1))/p(1)\}(1 + ka^{-1}\pi r^2))^{-1} \quad (41)$$

Combining (40) and (41) yields a quadratic equation in $p(1)$ for which there are two roots: 1 and the estimator

$$\hat{p}(1) = [1 + 1/(k\hat{a}^{-1}\pi r^2)](1/n) \sum_{i=1}^{n} 1(\bar{y}_i \neq 0) \quad (42)$$

The probability of a wet day followed by a wet day, $q(1)$, can be estimated analogously:

$$\hat{q}(1) = \sum_{i=1}^{n-1} 1(\bar{y}_i \neq 0, \bar{y}_{i+1} \neq 0) \Big/ \sum_{i=1}^{n-1} 1(\bar{y}_i \neq 0) + P\{X(i) = 0 \mid \bar{Y}_i = 0\} \cdot \left[\sum_{i=1}^{n-1} 1(\bar{y}_i \neq 0, \bar{y}_{i+1} = 0) \Big/ \sum_{i=1}^{n-1} 1(\bar{y}_i \neq 0) \right] \quad (43)$$

The parameter $q(0)$ can be obtained from (6).

The estimation procedure is initiated by specifying an initial value of $\hat{p}(1)$. A reasonable initial estimator is the fraction of days with recorded rainfall, i.e., the estimator of (40) with the second term dropped.

5. NUMERICAL EXAMPLES

Daily precipitation data from the Potomac River basin are used in this section to examine suitability of model assumptions and to illustrate implementation of the method of moments estimation procedure.

TABLE 1. Correlation of Daily Rainfall Totals for Five Gages in the Potomac River Basin

	Gage 1	Gage 2	Gage 3	Gage 4	Gage 5
Gage 1	1	0.64 (44)	0.63 (24)	0.41 (60)	0.39 (65)
Gage 2	0.64 (44)	1	0.57 (45)	0.35 (75)	0.36 (75)
Gage 3	0.63 (24)	0.57 (45)	1	0.48 (48)	0.46 (50)
Gage 4	0.41 (60)	0.35 (75)	0.48 (48)	1	0.69 (10)
Gage 5	0.39 (65)	0.36 (75)	0.46 (50)	0.69 (10)	1

Distances between gages (in miles) are given in parentheses. One mile equals 1.609 km.

TABLE 2. Coefficient of Variation of Storm Depth Totals for All Days With Recorded Rainfall and Days With Recorded Rainfall Greater than 0.10 Inches

Gage	Coefficient of Variation*	Coefficient of Variation†
1	1.30	0.97
2	1.45	1.04
3	1.49	1.14
4	1.39	0.97
5	1.37	0.97

One inch equals 2.54 cm.
*All days with recorded rainfall.
†Days with recorded rainfall greater than 0.10 inches.

The spatial range of applicability of the precipitation model is determined by the assumptions that (1) the entire region is under the influence of the same climatological state and (2) the region is climatologically homogeneous. Table 1 contains the correlation matrix (and in parentheses, intergage distances) of daily summer season (July to September) rainfall totals for five gages in the Potomac River basin. The period of record is 1961–1970 and each of the gages has an 8 A.M. observation time (these data sets are used for all of the analyses presented in this section). Gages 4 and 5, located on the Allegheny Plateau, are separated from gages 1, 2, and 3, which are located in the Valley and Ridge province, by a major climatological boundary, the Allegheny Front.

Note that while gages 3 and 4 and gages 3 and 2 are separated by comparable distances, the correlation is noticeably smaller for the first pair. In general, correlation between gages depends less on distance separating gages than on location of the gages relative to the Allegheny Front. Consequently, in developing a model for the Potomac River basin it would be necessary to separate the Allegheny Plateau portion from the Valley and Ridge portion.

One consequence of model assumptions (equation (23)) is that the spatial correlation function is constant for separation distances greater than the diameter of a rain cell. The results of Table 1 suggest that this assumption is not a gross oversimplification of the physical setting. Note, for example, that the correlation of the two Allegheny Front gages separated by less than ten miles is comparable to correlations between the Valley and Ridge gages separated by as much as 50 miles (80.45 km).

It was concluded from (25) that storm depths for days on which rainfall was recorded are exponentially distributed. Table 2 contains the coefficient of variation in storm depth totals for days with measurable precipitation at five gages in the Potomac River basin. The exponential assumption, under which the coefficient of variation is equal to 1, is not supported by this data. Table 2 also contains coefficient of variation of storm depth totals for days with greater than 0.1 inch (0.254 cm) of recorded rainfall. These results suggest it may be useful to adopt threshold precipitation levels larger than the reporting minimum of 0.01 inch (0.0254 cm).

The form of the "stochastic climatologicl process" is based on results presented in the work by *Smith and Karr* [1983]. There it is shown that summer season precipitation occurrences in the Potomac River basin can be represented by alternating "dry periods" dominated by anticyclonic con-

TABLE 3. Parameter Estimates Using Iterative Method of Moments Procedure

Iteration	a	b	$q(1)$	$p(1)$
1	5.5820	7.4741	0.6222	0.4869
2	9.6127	5.3041	0.6693	0.5508
3	12.5234	4.6056	0.7033	0.5969
4	14.6252	4.2741	0.7278	0.6302
5	16.1429	4.0884	0.7455	0.6543
10	19.3121	3.7947	0.7825	0.7045
20	20.0564	3.7392	0.7912	0.7163
30	20.0852	3.7371	0.7915	0.7168
40	20.0863	3.7371	0.7915	0.7168
50	20.0863	3.7371	0.7915	0.7168

ditions and "wet periods" during which air mass thunderstorms predominate.

The method of moments estimation procedure was applied to a 240-mi^2 (621 km^2) catchment (the North Branch of the Potomac River) located on the Allegheny Plateau. The region contains two rain gages separated by 10 miles (16.09 km). Based on results reported from Ohio for the Thunderstorm Project [*Byers and Braham*, 1949], cell radius r for the model was taken to be 2 miles (3.208 km).

Implementation of the iterative estimation procedure requires precomputing the statistics:

$$\hat{\mu} = (kn)^{-1} \sum_{i=1}^{k} \sum_{j=1}^{n} Y_j(i)$$

$$\hat{\sigma}^2 = (kn)^{-1} \sum_{i=1}^{k} \sum_{j=1}^{n} [Y_j(i) - \hat{\mu}]^2$$

$$\bar{p} = n^{-1} \sum_{i=1}^{n} 1(\bar{y}_i \neq 0)$$

$$\bar{q} = \sum_{i=1}^{n-1} 1(\bar{y}_i \neq 0, \bar{y}_{i+1} \neq 0) \Big/ \sum_{i=1}^{n-1} 1(\bar{y}_i \neq 0)$$

where k is the number of gages, and n is the number of days. Values of the sample statistics for the North Branch data are $\hat{\mu} = 0.12$, $\hat{\sigma}^2 = 0.09$, $\bar{p} = 0.40$, and $\bar{q} = 0.56$.

The iterative procedure commences with (38) using $\hat{p}(1) = \bar{p}$. Table 3 summarizes estimates for 50 iterations. The percent change in parameter estimates is less than 0.001 after 50 iterations. The estimated parameters are $\hat{a} = 20.1$, $\hat{b} = 3.7$, $\hat{p}(1) = 0.72$, and $\hat{q}(1) = 0.79$. From (6) we obtain $\hat{q}(0) = 0.46$.

The estimated parameters yield the following interpretations of model components. The percent of days during which the climatological state is wet is 0.72; on a wet day the probability that tomorrow will also be a wet day is 0.79. The mean number of rain cells (of radius 2 miles (3.208 km)) on a wet day is 0.05 per square mile, or 12 for the 240-mi^2 (621 km^2) basin. The mean rainfall intensity for a rain cell is 0.27 inches (0.69 cm).

6. Summary and Conclusions

In this paper we have presented a simple, physically based model of space-time rainfall and developed maximum likelihood and method of moments estimation procedures for data from a network of rain gages. It was noted in the Introduction that a primary goal of the paper is to examine the role of statistical parameter estimation in development of physically based precipitation models. We summarize below the main conclusions.

1. Estimation of a "cell radius" parameter from an operational network of rain gages will be difficult or impossible. The problems described in this paper for circular rain cells of fixed radius will be more acute for models which accomodate additional structure such as elliptical rain cells or cells with varying radii.

2. While both estimation procedures can be implemented, the method of moments procedure requires substantially less computational effort than the maximum likelihood procedure. The contrast in computational requirements increases with added model complexity; the maximum likelihood procedure becomes intractable under formulations only slightly more general than presented in this paper.

3. An important feature of both estimation procedures is the interaction between spatial and temporal parameters. In each there is a natural separation of spatial and temporal estimation components and a "link" between the spatial and temporal components.

4. Numerical results emphasize the role of spatial scale in model development. *Waymire et al.* [1984] have examined the role of spatial scale of meteorological systems for space-time rainfall modeling. Topographically induced inhomogeneities should also play a major role in determining the appropriate scale or formulation of space-time rainfall models.

5. A major issue that has not been addressed in this paper is properties of estimators. Further research should be undertaken to determine whether properties such as consistency, asymptotic normality, and efficiency can be established.

Appendix

The proof of (18) is sketched below. Two additional definitions are required. Let M be a point process on a Euclidean space E.

Definition

M is a Cox process directed by the random measure Λ if the following are true.

1. For every finite collection of disjoint sets $A_1, \cdots, A_n$, the random variables, $M(A_1), \cdots, M(A_n)$ are conditionally independent given Λ.

2. For all nonnegative integers k and all sets A,

$$P\{M(A) = k \mid \Lambda\} = \exp(-\Lambda(A))\Lambda(A)^k/k!$$

that is, the conditional distribution of $M(A)$ given Λ is Poisson with parameter $\Lambda(A)$.

Definition

The Laplace functional L_M of M is defined by

$$L_M(f) = E\left[\exp\left(-\int_E f(x)\, dM(x)\right)\right]$$

where f is a nonnegative function on E.

The following lemma [*Kallenberg*, 1976] is needed for the proof.

Lemma

If M is a Cox process directed by Λ, then the Laplace functional of M is given by

$$L_M(f) = L_\Lambda(1 - e^{-f})$$

where L_Λ is the Laplace functional of Λ.

Equation (16), which represents total rainfall on day n, can be rewritten as follows:

$$Z_n(A) = \int_R um(A - x)M_n(dx, du)$$

where M_n is a Cox process on $Rx[0, \infty)$. $M_n(A, B)$ can be interpreted as the number of rain cells located in the region A with maximum intensity falling in the set B. The directing measure of M_n is

$$\Lambda_n(A, B) = V_n \int_{A \times B} f(u)\, dx\, du$$

where f is the exponential density with parameter b. Thus we have

$$\begin{aligned}
&E[\exp(-sZ_n(A))] \\
&= E[\exp(-sZ_n(A))1(X(n) = 1)]P\{X(n) = 1\} + P\{X(n) = 0\} \\
&= p(1)E\left[\exp\left(-s\int_R um(A - x)M_n(dx, du)\right)\right] + (1 - p(1)) \\
&= p(1)L_{M_n}(usm(A - x)) + (1 - p(1))
\end{aligned}$$

The result follows from application of the lemma.

REFERENCES

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

Bras, R., and I. Rodriguez-Iturbe, Rainfall generation: A nonstationary time varying multidimensional model, *Water Resour. Res.*, *12*(3), 450–456, 1976.

Byers, H. R., and R. R. Braham, Jr., *The Thunderstorm*, U.S. Weather Bureau, U.S. Government Printing Office, Washington, D. C., 1949.

Cinlar, E., *Introduction to Stochastic Processes*, Prentice-Hall, Englewood Cliffs, N. J., 1975.

Corotis, R. B., Stochastic considerations in thunderstorm modeling, *J. Hydraul. Div. Am. Soc. Civ. Eng.*, *HY7*, 865–878, 1976.

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

Georgakakos, K. P., and R. L. Bras, A hydrologically useful station precipitation model, 2, Case studies, *Water Resour. Res.*, *20*(11), 1597–1610, 1984.

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

Kallenberg, O., *Random Measures*, Akademie Verlag, Berlin, 1976.

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.

Matern, B., *Spatial Variation*, Meddelanden fran Statens Skogsforstinstitut, Stockholm, 1960.

Rodhe, H., and J. Grandell, Estimates of characteristic times for precipitation scavaging, *J. Atmos. Sci.*, *38*(2), 370–386, 1981.

Rodriguez-Iturbe, I., and H. M. Mejia, Design of rainfall networks in space and time, *Water Resour. Res.*, *10*(4), 713–728, 1974.

Smith, J. A., and A. F. Karr, A point process model of summer season rainfall occurrences, *Water Resour. Res.*, *19*(1), 95–103, 1983.

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

A. F. Karr, Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, MD 21218.

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

(Received October 1, 1984;)
revised March 20, 1985;
accepted April 4, 1985.)