WATER RESOURCES RESEARCH, VOL. 23, NO. 5, PAGES 875–884, MAY 1987

A Markov Renewal Model for Rainfall Occurrences

EFI FOUFOULA-GEORGIOU

Department of Civil Engineering, Iowa State University, Ames

DENNIS P. LETTENMAIER

Department of Civil Engineering, University of Washington, Seattle

A probabilistic model for the temporal description of daily rainfall occurrences at a single location is presented. By defining an event as a day with measurable precipitation the model is cast into the discrete-time point process framework. In the proposed model the sequence of times between events is formed by sampling from two geometric distributions, according to transition probabilities specified by a Markov chain. The model belongs to the class of Markov renewal processes and exhibits clustering relative to the independent Bernoulli process. As a special case, it reduces to a renewal model with a mixture distribution for the interarrival times. The rainfall occurrence model coupled with a mixed exponential distribution for the nonzero daily rainfall amounts was applied to the daily rainfall series for Snoqualmie Falls, Washington, and was successful in preserving the short-term structure of the occurrence process, as well as the distributional properties of the seasonal rainfall amounts.

1. INTRODUCTION

Point process theory has been widely used to model the stochastic structure of short-term rainfall [cf. *Kavvas and Delleur*, 1981; *Gupta and Waymire*, 1979; *Waymire and Gupta*, 1981*a*, *b*; *Waymire et al.*, 1984; *Smith and Karr*, 1983, 1985]. Rainfall data are usually available in the form of cumulative amounts over disjoint equispaced time intervals. In adapting the continuous-time point process theory [cf. *Çinlar*, 1975; *Cox and Lewis*, 1978; *Cox and Isham*, 1980] to modeling short-term rainfall, two approaches can be followed. The first is to define an "event" as a day with measurable precipitation and develop discrete-time point process models to describe the probabilistic structure of the sequence of rainy and dry days. Such a probabilistic model is proposed in this paper. The second approach is to assume the existence of an underlying continuous-time rainfall occurrence process whose outcome is only observed as the integral of the continuous process over the given sampling interval. Under the second approach, one tries to infer the properties of the underlying continuous-time process from the observed discrete data. Results in this direction have been reported by *Rodriguez-Iturbe et al.* [1984], *Valdes et al.* [1985], and *Foufoula-Georgiou and Guttorp* [1986]. The main conclusion of these studies is that the inferred description of the underlying process depends on the time scale at which the fitting of the model is made. This poses limitations on the model in terms of inability to extrapolate at other time scales and inability to infer properties of the underlying rainfall-generating mechanism based on the sampled realizations. In addition, estimation problems arise when observations over relatively long sampling intervals, such as days, are used to estimate the parameters of continuous-time models [*Foufoula-Georgiou and Guttorp*, 1986].

The daily rainfall occurrence process has been extensively studied over the past two decades. The only discrete-time models investigated to date are Markov chains (see, for example, *Gabriel and Neumann* [1962]), the discrete autoregressive moving average models (DARMA) [*Chang et al.*, 1984], and a discrete-time alternating renewal model of [*Galloy et al.*, 1981]. Markov chains have been found, in general, to be inadequate to model the clustering dependencies present in daily rainfall occurrences. Models from the DARMA family [*Jacobs and Lewis*, 1978] were used by *Chang et al.* [1984], who reported satisfactory results in modeling daily rainfall occurrences in Indiana. In our view the main disadvantage of the DARMA models is the lack of physical motivation for the model structure and the discontinuous memory they exhibit [cf. *Keenan*, 1980]. On the other hand, point process theory permits more elegant mathematical formulations of intuitively appealing dependence properties of the process, such as the conditional intensity function or the index of dispersion, which also provide measures of clustering. Recently, *Smith* [1987] has proposed a new family of discrete point process models for daily rainfall occurrences. Theoretical and empirical comparisons of those models (termed Markov Bernoulli models) with the class of models proposed herein would be worth investigating.

Daily rainfall occurrences are the result of the interaction of several rainfall-generating mechanisms. For example, the first rainy day in a wet period may be the result of a frontal storm passing over a region, whereas subsequent rainy days in the same wet period may be considered secondary events. In that sense, times between events may come from different probability distributions, for instance, one with a small mean and coefficient of variation for the secondary events and one with a large mean and coefficient of variation for the primary events. The sequence of event types is governed by transition probabilities with higher probabilities of having secondary events after a primary event or after a small number of secondary events. In the model proposed in this paper the times between daily rainfall occurrences are sampled from two different probability distributions (this is called a two-state process) which we assume to be geometric. The transition from one interarrival type (or event type) to the other is governed by a Markov chain. This model belongs to a class known as Markov renewal models [cf. *Çinlar*, 1975; *Cox and Lewis*, 1978]. Markov renewal processes are, in general, nonrenewal, a nomenclature

Paper number 6W4795.
0043-1397/87/006W-4795$05.00

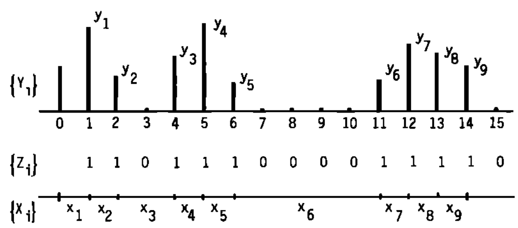

Fig. 1. Schematic representation of a daily rainfall process and definition of the $\{Y_i\}$, $\{Z_i\}$, and $\{X_i\}$ series. Note that the process starts at an arbitrary event. $\{Y_i\}$ is the series of the nonzero daily rainfall amounts, $\{Z_i\}$ is the binary series of rain-no rain, and $\{X_i\}$ is the series of interarrival times.

that may at first seem contradictory. The term Markov renewal refers to the conditional dependence of the present state (interarrival type) on the previous state only and not on states before that. By contrast, for a renewal process (which results as a special case of the Markov renewal process) the present state is independent of all previous states.

Apart from the intuitive appeal that discrete rainfall amounts represent the combined effect of several underlying mechanisms, a justification for using a mixture model may be provided by the form of the cumulative distribution function $F(x)$ of the interarrival times. For a geometric distribution with parameter p the log-survivor function ($\ln(1 - F(x))$) of the intervals is a straight line with slope $\ln(1 - p)$. Log-survivor functions of times between events at several stations located throughout the United States [*Foufoula-Georgiou and Lettenmaier*, 1986] suggest that the interarrival times come from two different geometric distributions. It should be noted, however, that graphical identification of mixtures is extremely difficult, in general, [cf. *Leytham*, 1984] and becomes even harder for discrete data.

The general class of Markov renewal processes, to which the proposed discrete-time point process model proposed belongs, were introduced by *Smith* [1955] and were later studied by *Pyke* [1961*a*, *b*] and *Cox* [1963]. An extensive bibliography of theoretical developments and applications of Markov renewal processes is given by *Teugels* [1976]. Markov renewal processes have a flexible dependence structure. It will be seen later that Markov chains, Markov processes, renewal processes, and alternating renewal processes [cf. *Çinlar*, 1975] are all special cases of the general Markov renewal process. In order to illustrate how our model differs from a Markov chain we briefly note that the probability of having a rainy day does not depend on the condition (rain-no rain) of the previous day but rather on the number of days since the last rain. Within a rainy period (consecutive rainy days), however, our process behaves as a Markov chain.

The emphasis of the work presented in this paper is on the modeling of the occurrence process. Others [e.g., *Woolhiser and Roldan*, 1982] have addressed the modeling of event scale rainfall amounts. Although this paper is concluded with an example in which both the occurrences and amounts are modeled, the amounts modeling part closely follows the work of others. It is included primarily for completeness and to allow assessment of the performance of the occurrence model as it affects the modeling of cumulative (seasonal) amounts.

2. Rainfall Occurrence Model

Consider the daily rainfall process schematically presented in Figure 1. Let $\{Y_i\}$ denote the series of nonzero daily rainfall amounts, $\{X_i\}$ denote the series of times between events (interarrival times), and $\{Z_i\}$ denote the binary series of zeros and ones, zeros for dry days and ones for wet days. Rainfall modeling at the event scale is best performed in two steps: the occurrence of rainfall is modeled first, followed by the modeling of the amounts; finally, the two models are superimposed.

Our rainfall occurrence model describes the sequence of interarrival times $\{X_i\}$. Useful properties of the binary series $\{Z_i\}$, such as transition probabilities between rainy and dry days, are subsequently derived. The model of interarrival times is a discrete-time Markov renewal model. A formal definition of a Markov renewal process in continuous time is given by *Çinlar* [1975, p. 313]:

Definition. For each $n \in N$, let a random variable S_n take on values in a countable set of states $E = \{1, 2, \cdots\}$ and a random variable T_n take on values in $R_+ = [0, +\infty)$ such that $0 = T_0 \le T_1 \le T_2 \le \cdots$. The stochastic process $(S, T) = \{S_n, T_n\}$, $n \in N$ is said to be a Markov renewal process with state space E provided that

$$P\{S_{n+1} = j, T_{n+1} - T_n \le t | S_0, \cdots, S_n; T_0, \cdots, T_n\} = P\{S_{n+1} = j, T_{n+1} - T_n \le | S_n\} \quad (1)$$

for all $n \in N, j \in E$, and $t \in R_+$.

For the rainfall occurrence model the random variable S_n is given the interpretation of the "type" (or "state") of an interarrival time and takes on values from the binary set $E = \{1, 2\}$. This is a two-state Markov renewal model where the two types of interarrival times (type 1 and type 2) are sampled according to a Markov chain with state space E. Let $\langle X_i \rangle$ denote the type of the ith interarrival time, that is, $\langle X_i \rangle = 1, 2$ for type 1 and type 2, respectively. The transition probability matrix of the Markov chain is

$$P = \begin{bmatrix} a_1 & 1 - a_1 \\ 1 - a_2 & a_2 \end{bmatrix} \quad (2)$$

where

$$a_j = P\{\langle X_i \rangle = j | \langle X_{i-1} \rangle = j\} \qquad j = 1, 2 \quad (3)$$

For example, given that the interarrival time X_{i-1} is of the type 1, the probability that X_i will also be of type 1 is a_1. Associated with the Markov chain are the limit or equilibrium probabilities

$$e_j = \lim_{i \to \infty} P\{\langle X_i \rangle = j\} \qquad j = 1, 2 \quad (4)$$

which are the unconditional probabilities of any interval X_i being of type 1 or type 2. Note that $e_2 = 1 - e_1$. From the theory of Markov chains [cf. *Cox and Miller*, 1965] it is known that

$$e_1 = \frac{1 - a_2}{2 - a_1 - a_2} \quad (5)$$

Note that if the conditional probabilities a_1 and a_2 are equal to the unconditional probabilities e_1 and e_2 (in that case, $a_1 + a_2 = 1$), the process of the types of interarrival times reduces to a renewal process.

To complete the model description, we need to specify what the type 1 and type 2 interarrival times mean. An interarrival time X_i is said to be of type 1 (or type 2) if it is sampled from a probability distribution $f_1(x)$ (or $f_2(x)$). For the rainfall occurrence model these distributions have been assumed geometric with parameters p_1 and p_2, respectively. One can write therefore that

$$f_j(x_i) = P\{X_i = x_i | \langle X_i \rangle = j\} = p_j(1-p_j)^{x_i - 1} \qquad j = 1, 2 \tag{6}$$

The assumption of geometric distributions is supported by the data analysis presented later in this paper and by *Foufoula-Georgiou* [1985]. In the rest of this section the statistical properties of intervals and counts for the proposed two-state Markov renewal model are derived.

2.1. *Interval Properties*

The moment-generating function of the interarrival times of a two-state Markov renewal model is given as

$$\psi(z) = e_1\psi_1(z) + e_2\psi_2(z) \tag{7}$$

where $\psi_j(z)$, $j = 1, 2$ is the moment generating function of the probability distribution of the type j intervals. For a geometric distribution with parameter p,

$$\psi(z) = \frac{pz}{1 - (1-p)z} \tag{8}$$

[cf. *Parzen*, 1960]. Moments of the interarrival times are then obtained from

$$E(X^k) = (-1)^k \left. \frac{d^k\psi(z)}{dz^k} \right|_{z=1} \tag{9}$$

For instance, the mean, variance, and survivor function of the interarrival times are given by

$$E(X) = e_1/p_1 + e_2/p_2 \tag{10a}$$

$$\text{Var}(X) = e_1(1-p_1)/p_1^2 + e_2(1-p_2)/p_2^2 + e_1e_2(1/p_1 - 1/p_2)^2 \tag{10b}$$

$$R(x) = e_1(1-p_1)^x + e_2(1-p_2)^x \tag{10c}$$

It is important to note here that the proposed model admits coefficients of variation of interarrival times with values less or greater than one. In contrast, both the continuous-time Neyman-Scott [e.g., *Kavvas and Delleur*, 1981] and the doubly stochastic Poisson [*Smith and Karr*, 1983] processes have coefficients of variation always greater than one. In addition, it can be shown after some algebra that the coefficient of variation of the proposed Markov renewal model is always greater than $1 - m$, where $m = 1/E(X)$ is the rate of occurrence of the process. This observation suggests that the proposed process is always overdispersed (more clustered) than an independent Bernoulli process with the same rate of occurrence, which would have a coefficient of variation equal to $1 - m$. This is a desirable property since an analysis of several rainfall series [*Foufoula-Georgiou and Lettenmaier*, 1986] suggests that most daily rainfall occurrence series are overdispersed relative to the Bernoulli process.

The autocorrelation function of the interarrival times of the two-state Markov renewal process takes the form [cf. *Cox and Lewis*, 1978, p. 196]

$$\rho_k = c\beta^k \tag{11}$$

where

$$c = \frac{e_1e_2(1/p_1 - 1/p_2)^2}{e_1(1-p_1)/p_1^2 + e_2(1-p_2)/p_2^2 + e_1e_2(1/p_1 - 1/p_2)^2} \tag{12}$$

$$\beta = a_1 + a_2 - 1 \tag{13}$$

Consequently, the spectral density function of the intervals is

$$f_+(\omega) = \frac{1}{\pi}\left(1 + 2c\,\frac{\beta\cos\omega - \beta^2}{1 + \beta^2 - 2\beta\cos\omega}\right) \qquad 0 \le \omega \le \pi \tag{14}$$

Note that the autocorrelation function of the intervals becomes zero (renewal process) for $a_1 + a_2 = 1$, in which case the Markov chain of the type of intervals has conditional probabilities of occurrence equal to the unconditional ones, that is, it reduces to a Bernoulli process. Without loss of generality, we can assume that the type 1 interarrival times are sampled from the geometric distribution with the smaller mean. Then, for persistent structures (clustering) the conditional probability of being in state 1 is greater than the unconditional probability of (5), resulting in $a_1 + a_2 > 1$. In this case, the interarrival times have a positive autocorrelation function decaying with a rate $(a_1 + a_2 - 1)$.

2.2. *Count Properties*

Let

$$Z_k = 1(Y_k > \varepsilon) \qquad k \ge 0 \tag{15}$$

be the binary series of zeros and ones, where $1(E)$ is an index function taking the value of 1 if E occurs and zero otherwise, and where zeros (ones) correspond to days with cumulative rainfall less (greater) than ε. The small quantity ε (consistent with the previous definition of an event) has been taken equal to 0.01 inches. In this section we compute the statistical properties of the Z_k series in terms of the four parameters a_1, a_2, p_1, and p_2 of the Markov renewal model. The rate of occurrence of the Z_k process is

$$m = P\{\cdots\} = P\{Z_k = 1\} = \frac{p_1p_2(2 - a_1 - a_2)}{p_1(1-a_1) + p_2(1-a_2)} \tag{16}$$

One of the most descriptive statistical properties of a continuous-time stationary point process is its conditional intensity function [cf. *Cox and Lewis*, 1978, p. 73]. An analogous property for a discrete-time point process may be defined as

$$h_k = P\{Z_{t+k} = 1 | Z_t = 1\} = P\{Z_k = 1 | Z_0 = 1\} \qquad k = 1, 2, \cdots \tag{17}$$

which essentially defines a sequence of conditional probabilities of occurrence (note that the last equality is due to stationarity). The interpretation of h_k with respect to clustering remains the same as in the continuous case; values of h_k greater than the constant (unconditional) probability of occurrence m imply that the chance of having an event at time $t + k$ due to an event at time t is greater than the chance of having an event at any arbitrary time. Below we give the expression for h_k.

Proposition. The conditional probability of occurrence h_k of the discrete-time two-state Markov renewal process described above takes the form

$$h_k = m + AW^{k-1} \qquad k = 1, 2, \cdots \tag{18}$$

where

$$A = e_1p_1 + e_2p_2 - m \tag{19}$$

$$W = 1 - p_1(1-a_1) - p_2(1-a_2) \tag{20}$$

An outline of the proof is given in Appendix A.

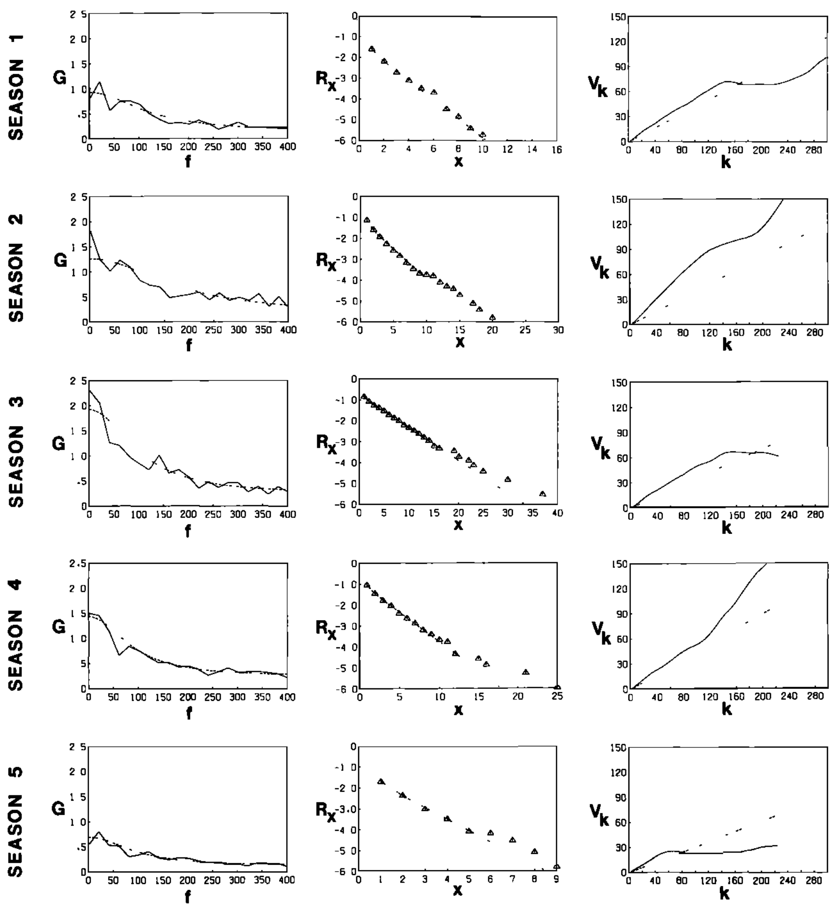

Fig. 2. Comparison of empirical (solid lines and open triangles) and theoretical (dashed lines) functions of the Markov renewal model fitted to daily rainfall occurrences at Snoqualmie Falls, Washington. G versus f is the normalized spectrum of counts $g_+'(\omega)$ versus frequency factor $f = \omega T/2\pi$, R_x versus x is the log-survivor function $\ln R(x)$ versus interarrival time x (days), and V_k versus k is the variance of counts V_k versus interval length k (days). For more details on these functions see text.

It is particularly important to note in (18) that since A is positive and $0 < W < 1$, the conditional intensity function decreases geometrically to the constant intensity m of the process. This implies that the Markov renewal process exhibits clustering. The shape of the conditional intensity function is only indicative of the presence, but not the type, of clustering. However, the fact that the coefficient of variation of the interarrival times is always greater than the coefficient of variation of the Bernoulli process suggests that the form of clustering is overdispersion relative to the Bernoulli process. This is the type of clustering found in most daily rainfall occurrence series [e.g., *Foufoula-Georgiou*, 1985].

Having an expression for the conditional intensity function, all the other properties of the counting process can be readily obtained. The expected number of events within a period of k time units (for example, days) after the occurrence of an event is given as

$$H_k = mk + A\sum_{i=1}^{k} W^i = mk + A\frac{W^{k+1}-1}{W-1} \qquad k = 1, 2, \cdots \tag{21}$$

Notice that as $k \to \infty$, $H_k - mk \to 0$, where mk is the expected number of events in any one period of length k time units. The variance of counts, that is, the variance of the number of events in a period of k time units after the occurrence of an event, is

$$V_k = mk - m^2k^2 + 2m\sum_{i=1}^{k-1}(k-i)h_i \tag{22}$$

where h_i is given by (18). Finally, the index of dispersion is $I_k = V_k/mk$. Note that the above formulae for V_k and I_k apply to any discrete-time point process with conditional intensity function h_k (see also *Guttorp* [1986]).

The spectrum of counts, $g_+(\omega)$, of the two-state Markov renewal process is

$$g_+(\omega) = \frac{m}{\pi}\left(1 - m - 2A\frac{W - \cos\omega}{1 - 2W\cos\omega + W^2}\right) \qquad 0 \le \omega \le \pi \tag{23}$$

and is computed by simply taking the Fourier transform of the covariance of counts $c_k = E(Z_l Z_{l+k}) = m(h_k - m) = mAW^{k-1}$. The normalized spectrum of counts is defined as $g_+'(\omega) = \pi g_+(\omega)/m$ and is usually plotted (see Figure 2) versus a frequency factor $f = \omega T/2\pi$, where T is the total length of observation.

3. METHODS FOR FITTING THE MARKOV RENEWAL MODEL

The discrete-time Markov renewal model developed in the previous section has four parameters: a_1, the transition probability from type 1 to type 1 interval; a_2, the transition probability from type 2 to type 2 interval; p_1, the parameter of the geometric distribution of the type 1 intervals; and p_2, the parameter of the geometric distribution of the type 2 intervals. Note that the interarrival times cannot be classified directly as belonging to type 1 or type 2 by observation of the series of daily rainfall events. Only probabilistic classification is possible. Thus the transition probabilities a_1 and a_2 must be estimated together with the parameters of the two geometric distributions p_1 and p_2. In the following section, maximum likelihood and method of moments estimators for the parameters a_1, a_2, p_1, and p_2 are studied.

3.1. *Maximum Likelihood Estimation*

The observations to which the Markov renewal model is fitted are the interarrival times X_i, that is, the sequence of lengths of dry periods between consecutive rainy days. Let $\boldsymbol{\theta} = (a_1, a_2, p_1, p_2)$ denote the vector of unknown parameters and $(x_1, x_2, \cdots, x_n)$ the sampled sequence of n interarrival times.

Proposition. The likelihood function of the two-state Markov renewal model takes the form

$$L(\boldsymbol{\theta}|x_1, \cdots, x_n) = EB_1PB_2P, \cdots, B_n' \tag{24}$$

where the matrices $E, B_1, \cdots, B_n'$ are functions of the four parameters of the model and the known sampled data:

$$E = (e_1 1 - e_1) = \left(\frac{1-a_2}{2-a_1-a_2}\ \frac{1-a_1}{2-a_1-a_2}\right) \tag{25a}$$

$$B_i = \begin{pmatrix} p_1(1-p_1)^{x_i-1} & 0 \\ 0 & p_2(1-p_2)^{x_i-1} \end{pmatrix} \quad i = 1, \cdots, n \tag{25b}$$

$$B_n' = B_n\begin{pmatrix}1\\1\end{pmatrix} \tag{25c}$$

and P is the transition probability matrix of the Markov chain as defined in (2). The proof of the above proposition is sketched in Appendix B.

Note that for independent interarrival times the process reduces to a renewal process with a mixed geometric distribution for the interarrival times. In this case, the log likelihood function is simply

$$L'(\boldsymbol{\theta}|x_1, \cdots, x_n) = \sum_{i=1}^{n} \ln\,[e_1p_1(1-p_1)^{x_i-1} + (1-e_1)p_2(1-p_2)^{x_i-1}] \tag{26}$$

3.2. *Method of Moments Estimation*

The method of moments (MOM) uses sample estimates of the first three moments and the lag 1 covariance of the interarrival times and solves for the four parameters by using the theoretical relationships between the population moments and the parameters. This method has a major drawback in that the estimate of the third moment is highly variable, so the resulting parameter estimates can be unstable. A modified method of moments estimation which uses the median instead of the third moment was also tested [*Foufoula-Georgiou*, 1985]. Due to the discreteness of the data, however, the median has poor sampling properties, which lead to unsatisfactory performance of this method. Therefore it was dropped from further consideration.

It should be noted that the moments of the interarrival times involve only the equilibrium unconditional probabilities e_1 and $e_2 = 1 - e_1$. The transition probabilities are introduced only in the second product moments, as, for example, in the autocorrelation coefficient r_1. Therefore it is possible to use the first three moments for estimation of e_1, p_1, and p_2 and then use the first autocorrelation coefficient $r_1 = c(a_1 + a_2 - 1)$, where c is given in (12), together with e_1 of (5) to solve for a_1 and a_2:

$$a_1 = (1-e_1)(r_1/c + 1) + 2e_1 - 1 \tag{27a}$$

$$a_2 = e_1(r_1/c + 1) - 2e_1 + 1 \tag{27b}$$

From the above two equations one can see that for acceptable parameter estimates, that is, $0 < a_1, a_2 < 1$, the following inequality must hold

$$-\min\,(e_1/e_2, e_2/e_1) < r_1/c < 1 \tag{28}$$

Note that the value $\min\,(e_1/e_2, e_2/e_1)$ corresponds to the ratio of the smallest to the largest equilibrium probability, a value always less than 1. Therefore inequality (28) is consistent with the requirement that the autocorrelation function of the process, given as $r_k = c(a_1 + a_1 - 1)^k$, is less than 1 in absolute value.

4. STATISTICAL PROPERTIES OF THE ESTIMATORS

The two methods discussed in the previous section were tested for consistency (bias) and efficiency (variability) using Monte Carlo simulation. Several sets of population parameters were selected to represent a range of underlying processes consistent with the data analysis reported by *Foufoula-Georgiou and Lettenmaier* [1986]. Two kinds of dependencies were considered in selecting population parameters: dependency in the intervals (a measure of which is the autocorrelation function r_k) and dependency in the counts or clustering (a measure of which is the conditional intensity function h_k). The type of clustering (overdispersion and underdispersion relative to the Bernoulli process) is further inferred by the variance time curve and index of dispersion. It should be emphasized that independence in intervals does not imply or result from independence in counts. For instance, a renewal process may well be clustered as, for example, the renewal Cox process with Markovian intensity [*Smith and Karr*, 1983] and the renewal form of the Markov renewal process discussed herein. In the discussion that follows the dependencies in both the intervals and counts are used to characterize the underlying process. Recall that, for a Markov renewal process, these dependencies take the form $r_k = c(a_1 + a_2 - 1)^k$ and $h_k = m + AW^{k-1}$, with A and W defined in (19) and (20).

The first set of parameters tested was $\{a_1 = 0.4, a_2 = 0.3, p_1 = 0.8, p_2 = 0.2\}$. These parameter values correspond to an occurrence process with a mean interarrival time of 2.98 days, a standard deviation of 3.59 days (coefficient of variation c_v = 1.2), a skewness coefficient $c_s = 3.01$, and a first autocorrelation coefficient $r_1 = 0.08$. The conditional intensity function is $h_k = 0.335 + 0.186(0.38)^{k-1}$, which indicates a clustering of counts. Five hundred synthetic sequences of 50, 100, 200, 500, and 800 events (corresponding to approximately 150, 300, 600, 1500, and 2400 days of observation, as inferred by the rate of occurrence $m = 0.34$ events per day) were generated from a Markov renewal model with the above parameters. ML and MOM parameter estimates were computed for all synthetic sequences. (For the maximization of the likelihood

TABLE 1. Bias and Root-Mean-Square Error of the Parameter Estimates of a Markov Renewal Model With True Parameters $a_1 = 0.4$, $a_2 = 0.3$, $p_1 = 0.8$, and $p_2 = 0.2$

				Bias				RMSE			
N	m	m'	Method	a_1	a_2	p_1	p_2	a_1	a_2	p_1	p_2
50	500	500	ML	−0.0076	−0.0100	0.0224	0.0063	0.2307	0.2220	0.1322	0.0537
		159	MOM	0.1134	0.0176	−0.0044	−0.0144	0.2955	0.2339	0.2062	0.0692
100	500	500	ML	−0.0052	−0.0027	0.0093	0.0023	0.1711	0.1637	0.1050	0.0392
		238	MOM	0.0433	0.0385	0.0182	−0.0049	0.2543	0.2400	0.1847	0.0511
200	500	500	ML	−0.0098	−0.0045	0.0087	0.0007	0.1216	0.1177	0.0775	0.0267
		291	MOM	0.0317	0.0053	−0.0091	−0.0090	0.2175	0.1777	0.1847	0.0412
500	500	500	ML	0.0019	−0.0042	0.0014	−0.0003	0.0464	0.0169	0.0786	0.0710
		322	MOM	0.0197	−0.0085	0.0012	−0.0057	0.1754	0.1314	0.1544	0.0314
800	500	500	ML	−0.0017	0.0018	0.0012	0.0004	0.0640	0.0578	0.0384	0.0135
		212	MOM	0.0090	−0.0075	0.0028	−0.0056	0.1612	0.1307	0.1500	0.0291

N is the number of events in each sequence, m is the number of sequences, and m' is the number of sequences a method succeeded. ML is the maximum likelihood, MOM is the method of moments, and RMSE is the root-mean-square error.

function the simplex method described by *Nelder and Mead* [1965] was used.) The bias and root-mean-square error of both parameter estimators are shown in Table 1. As expected, the consistency (bias) and efficiency (variability) of the estimators improve as the number of events increases. The ML is clearly superior to the MOM, although even the latter performs satisfactorily for moderately large samples (500 events). One major drawback of the MOM is that it often failed to obtain feasible parameter estimates. This is partly due to failure of the iterative scheme to converge within the specified criteria and also due to infeasibility (equation (28)) of the obtained parameter estimates due to the large variability of r_1 in small samples. Refinement of the algorithm would probably decrease the number of failures.

The second set of parameters tested was $\{a_1 = 0.9, a_2 = 0.6, p_1 = 0.8, p_2 = 0.4\}$. These parameters correspond to an occurrence process with mean interarrival time 1.5 days ($m = 0.667$), a standard deviation of 1.11 days ($c_v = 0.74$), and a skewness coefficient $c_s = 4.02$. The autocorrelation function of the process is $r_k = 0.1(0.5)^k$, and the conditional intensity function is $h_k = 0.667 + 0.05(0.76)^{k-1}$. These functions indicate a strong dependence structure in the intervals but a relatively small clustering in the counts. These properties together with the small mean and variance of the lengths of interarrival times make this process difficult to identify. This expectation is confirmed by the results of Table 2, which shows that the ML method starts performing satisfactorily only for sample sizes greater than 200. Fortunately, for the rainfall series, small and less variable interarrival times are always associated with larger sample sizes (see, for example, Table 4).

The effect of the dependence in the intervals (as measured by the first autocorrelation coefficient of the process) on the consistency and efficiency of the estimators a_1, a_2, p_1 and p_2 was also tested. For the discussion that follows, the convention is made that e_1 corresponds to the geometric distribution with the larger parameter (the distribution with the shorter tail). A value of $e_1 > 0.5$ therefore implies that shorter interarrival times have greater probability of occurrence. The parameter set considered is $\{e_1 = 0.6, p_1 = 0.9, p_2 = 0.1\}$. Depending on the value of the first autocorrelation coefficient r_1, several sets of transition probabilities (a_1, a_2) were selected and the corresponding processes tested. Table 3 shows the results of this experiment. The ML estimator performs consistently well, while the performance of MOM becomes much poorer as the population parameters approach those of a renewal process.

The Monte Carlo results show conclusively that the ML method is much superior. The results also provide confidence that ML estimation will result in unbiased and consistent estimators for sample sizes that are commonly available in practice.

TABLE 2. Bias and Root-Mean-Square Error of the Parameter Estimates of a Markov Renewal Model With True Parameters $a_1 = 0.9$, $a_2 = 0.6$, $p_1 = 0.8$, and $p_2 = 0.4$

				Bias				RMSE			
N	m	m'	Method	a_1	a_2	p_1	p_2	a_1	a_2	p_1	p_2
50	500	500	ML	−0.1931	−0.0337	0.0650	0.0923	0.3313	0.3325	0.1257	0.1950
		180	MOM	−0.1158	−0.1837	0.0448	−0.1292	0.2578	0.3202	0.1470	0.2377
100	500	500	ML	−0.1182	−0.0369	0.0546	0.0482	0.2592	0.3031	0.1086	0.1493
		258	MOM	−0.0974	−0.1357	0.0437	−0.0650	0.2278	0.3006	0.1344	0.1875
200	500	500	ML	−0.0628	−0.0379	0.0279	0.0181	0.1833	0.2589	0.0825	0.1137
		321	MOM	−0.0785	−0.0913	0.0453	−0.0247	0.1755	0.2693	0.1058	0.1459
500	500	500	ML	−0.0298	−0.0184	0.0157	0.0065	0.1082	0.1729	0.0513	0.0776
		423	MOM	−0.0379	−0.0092	0.0338	0.0096	0.1094	0.2200	0.0737	0.0974

N is the number of events in each sequence, m is the number of sequences, and m' is the number of sequences a method succeeded. ML is the maximum likelihood, MOM is the method of moments, and RMSE is the root-mean-square error.

TABLE 3. Bias and Root-Mean-Square Error of the Parameter Estimates of Markov Renewal Models With Various Sets of True Parameters (a_1, a_2) Consistent With the Fixed Parameters $e_1 = 0.6$, $p_1 = 0.9$, and $p_2 = 0.1$

					Bias				RMSE			
(a_1, a_2) r_1	N	m	m'	Method	a_1	a_2	p_1	p_2	a_1	a_2	p_1	p_2
(0.4, 0.1)	200	500	500	ML	−0.0006	−0.0002	−0.0006	0.0011	0.0613	0.0464	0.0339	0.0122
−0.17	200		123	MOM	−0.0323	0.0378	−0.0279	0.0038	0.1818	0.1049	0.1418	0.0181
(0.52, 0.28)	200	500	500	ML	0.0059	−0.0014	−0.0010	0.0011	0.0632	0.0698	0.0415	0.0121
−0.068	200		210	MOM	0.0233	0.0122	−0.0888	−0.0039	0.1961	0.1794	0.2343	0.0445
(0.6, 0.4)	200	500	500	ML	−0.0002	−0.0082	0.0006	0.0018	0.0608	0.0783	0.0431	0.0130
0.0	200		290	MOM	0.0275	−0.0072	−0.1657	0.0019	0.2169	0.2072	0.3348	0.0427
(0.8, 0.7)	200	500	500	ML	−0.0002	−0.0076	0.0021	0.0020	0.0458	0.0664	0.0376	0.0127
+0.17	200		279	MOM	0.0358	−0.0553	−0.1145	0.0010	0.1130	0.2014	0.2900	0.0499

N is the number of events in each sequence, m is the number of sequences, and m' is the number of sequences a method succeeded. ML is the maximum likelihood, MOM is the method of moments, and RMSE is the root-mean-square error. The autocorrelation coefficient is indicated as r_1.

5. ANALYSIS OF DAILY RAINFALL DATA

Fifteen years of daily rainfall data from Snoqualmie Falls, Washington, were analyzed and subsequently modeled by the Markov renewal model described in the previous sections. One important aspect of modeling a periodic process such as daily rainfall is the selection of seasons within which the process can be reasonably assumed stationary. Our approach to selection of seasons consisted of two steps. First, a complete statistical analysis of the daily rainfall occurrence process as well as of the nonzero daily rainfall amounts was performed on a monthly basis. Then, after careful examination of the qualitative and quantitative similarities of the statistical properties of the monthly rainfall counts and amounts, months were grouped together and a seasonal analysis was subsequently performed. The seasonal properties were then checked for agreement with the monthly properties for all the months in each season. Following such a seasonal discrimination procedure, the seasons identified for Snoqualmie Falls were: season 1: January, February, and March; season 2: April, May, and June; season 3: July and August; season 4: September and October; and season 5: November and December. For these seasons a statistical analysis of intervals and counts was performed. Table 4 shows the statistics of the interarrival times, and Table 5 lists the first five autocorrelation coefficients. An approximate test for the hypothesis of significant autocorrelation of intervals results from assuming that the estimated autocorrelation coefficient r_j is distributed as $N(0, 1/(n-j)^{1/2})$. Using this test, the hypothesis that the interarrival times for all seasons are different than zero was rejected at the 95% confidence level. This test, although approximate, suggests that a renewal model with a mixed geometric distribution for the interarrival times would probably suffice for this data set. We have nonetheless elected to fit the Markov renewal model primarily for illustrative purposes. While fitting of the more parameter-parsimonious renewal process might be preferred in this special case, the Monte Carlo results (section 4) show that when ML parameter estimates are used the Markov renewal model can be successfully fit even when the true process is close to a renewal process. Therefore modest overfitting of the model is not a major concern.

Table 6 shows the maximum likelihood and method of moments parameter estimates of the Markov renewal model. For most seasons, $a_1 + a_2 \simeq 1$, confirming that the modeled process is very nearly a renewal process. The comparison of the empirical normalized spectrum of counts, log-survivor function, and variance time curve with their theoretical counterparts is shown in Figure 2. It is seen that the spectra of counts and the log-survivor functions are well preserved for all seasons. Although the theoretical variance time curves deviate from the empirical ones, they are, in fact, not significantly different given the wide confidence intervals of this highly variable function [cf. *Cox and Lewis*, 1978, p. 116].

The statistical properties of the nonzero daily rainfall amounts are shown in Tables 7 and 8. On the basis of previous research [*Woolhiser and Roldan*, 1982] three marginal distributions (Weibull, Gamma, and mixed exponential) were fitted to the nonzero daily rainfall amounts. The mixed exponential distribution was found to give the best fit. The maximum likelihood parameter estimates for the five seasons are given in Table 9. As far as the dependence structure of the amounts is concerned, it was found that the first autocorrelation coefficient for seasons 1 and 5 (January–March and November–December) were significant (although small) at

TABLE 4. Statistics of the Interarrival Times

Season	$\bar{x}$	s_x	c_v	c_s	Number of Events
1	1.496	1.377	0.920	4.217	896
2	2.101	2.603	1.239	4.085	672
3	3.715	5.235	1.409	2.924	246
4	2.271	2.776	1.222	3.781	391
5	1.393	1.125	0.808	4.212	657

TABLE 5. Autocorrelation Coefficients of Interarrival Times

Autocorrelation Coefficient	Season 1	2	3	4	5
r_1	0.047	−0.054	−0.010	0.025	0.036
r_2	−0.026	0.074	0.057	−0.026	0.034
r_3	0.032	0.016	−0.022	0.032	−0.040
r_4	−0.022	−0.009	0.051	−0.053	−0.040
r_5	0.003	0.042	−0.050	−0.042	−0.031

TABLE 6. Maximum Likelihood Estimates of the Parameters of a Markov Renewal Model Fitted to Daily Rainfall Occurrences at Snoqualmie Falls, Washington

Season	a_1	a_2	p_1	p_2
1	0.759	0.340	0.960	0.365
2	0.616	0.289	0.913	0.252
3	0.509	0.405	0.933	0.145
4	0.614	0.416	0.915	0.247
5	0.721	0.256	0.969	0.425

TABLE 8. Autocorrelation Coefficients of Nonzero Daily Rainfall Amounts

Autocorrelation Coefficient	Season				
	1	2	3	4	5
r_1	0.238*	0.058	0.094	0.054	0.130*
r_2	0.008	0.016	0.011	0.016	0.023
r_3	−0.014	−0.007	−0.046	0.057	0.014
r_4	0.011	−0.026	−0.023	0.025	−0.003
r_5	0.051	−0.010	0.012	0.014	−0.009

*Significant at the 5% level.

the 5% level. We elected not to attempt to preserve these correlations in our modeling of the amounts. Although this could be done, for example, by drawing the amounts from a log-normal lag 1 Markov process, this would be accomplished by sacrificing the preservation of the third moment (skew coefficient) of the amounts. Another alternative would be to use Exponential-AR or Gamma-AR models (see, for example, *Lawrance* [1980] and *Gaver and Lewis* [1980]).

6. Preservation of the Cumulative Rainfall Amounts

For the purposes of streamflow prediction or other applications where a mass balance is desired one is interested in the distribution of the total rainfall over the next t days. For example, for rainfall/runoff studies an important property of a daily rainfall generation scheme is its ability to preserve the total rainfall amounts over accounting periods such as a month or a year. The statistical properties of the accumulated rainfall process are given below (see also *Smith and Karr* [1983]).

Let $P(t)$ denote the accumulated rainfall process over a period of length t, that is,

$$P(t) = \sum_{i=1}^{N_t} Y_i \quad (29)$$

where $\{Y_i\}$ is the process of the nonzero daily rainfall amounts and $\{N_t\}$ is the daily rainfall occurrence process. If the assumption is made that the nonzero daily rainfall amounts are independent and identically distributed and that they are independent of the daily rainfall occurrences, the mean and variance of $P(t)$ are given as

$$E[P(t)] = \mu_y mt \quad (30a)$$

$$\text{Var}\,[P(t)] = \mu_y^2 mt + \sigma_y V(t) \quad (30b)$$

where $\mu_y = E[Y_i]$, $\sigma_y^2 = \text{Var}\,(Y_i)$, $V(t)$ is the variance time curve of the counting process $\{N_t\}$, and m is its rate of occurrence. For a Markov renewal model, m and $V(t)$ are given in terms of the parameters a_1, a_2, p_1, and p_2 from (16) and (22), respectively. For a mixed exponential distribution, μ_y and σ_y^2 are given in terms of the parameters a, λ_1, λ_2 by

$$\mu_y = a/\lambda_1 + (1 - a)/\lambda_2 \quad (31a)$$

$$\sigma_y^2 = a/\lambda_1^2 + (1 - a)/\lambda_2^2 + a(1 - a)(1/\lambda_1 - 1/\lambda_2)^2 \quad (31b)$$

It is understood that since there is a (slight) dependence in the Snoqualmie Falls rainfall amounts, (30a) and (30b) provide only approximations to the properties of the cumulative rainfall amounts. Due to averaging, however, they are expected to provide good approximations for cumulative rainfall amounts over long periods (for example, months) but not as good for shorter periods (for example, days). From Table 10 it can be seen that the derived properties of the Snoqualmie Falls cumulative seasonal rainfall amounts are in very good agreement with their empirical counterparts and are worse for seasons 1 and 5 in which the assumption of independence is least valid.

7. Summary and Conclusions

Point process theory provides a powerful tool for modeling the clustering present in rainfall. However, almost all of the available point process models are continuous in time and are not directly applicable to discretely sampled data such as the occurrence of daily rainfall. In this paper an alternative discrete-time point process model applicable to daily rainfall occurrences was introduced and its statistical properties derived. The model belongs to the class of Markov renewal precesses and is, in general, a nonrenewal clustered (overdispersed relative to the independent Bernoulli) process. Its flexible dependence structure is able to reproduce the types of clustering found in the daily rainfall occurrence structures analyzed by *Foufoula-Georgiou* [1985].

In the proposed model the sequence of times between events is formed through sampling from two geometric distributions according to transition probabilities specified by a Markov chain. As a special case, the proposed model includes a renewal process with a mixture distribution for the interarrival times. Methods of moments and maximum likelihood were

TABLE 7. Statistics of the Nonzero Daily Rainfall Amounts

Season	$\bar{x}$	s_x	c_v	c_s
1	0.373	0.456	1.222	3.019
2	0.240	0.281	1.170	2.559
3	0.216	0.274	1.270	2.116
4	0.311	0.342	1.099	1.963
5	0.407	0.474	1.164	2.165

TABLE 9. Maximum Likelihood Estimates of the Parameters of a Mixed Exponential Distribution Fitted to the Nonzero Daily Rainfall Amounts

Season	α	λ_1	λ_2
1	0.182	17.627	2.257
2	0.201	17.033	3.504
3	0.412	17.500	3.065
4	0.120	26.743	2.855
5	0.152	19.654	2.123

TABLE 10. Comparison of the Empirical and Theoretical Rainfall Seasonal Means and Standard Deviations

	Mean		Standard Deviation	
Season	Empirical	Theoretical	Empirical	Theoretical
1	22.145	22.321	5.821	4.709
2	10.621	10.386	2.643	2.609
3	3.352	3.595	1.587	1.651
4	8.689	8.400	2.749	2.782
5	17.789	17.833	4.445	4.226

Measurements in inches.

presented and the properties of the respective estimators studied via Monte Carlo simulation.

The Markov renewal model was fitted to daily rainfall occurrences for five seasons at Snoqualmie Falls, Washington. The adequacy of the model fit was confirmed by comparing the empirical normalized spectrum of counts, log-survivor function, and variance time curve, with their fitted counterparts. The Markov renewal model coupled with a mixed exponential distribution for the nonzero daily rainfall amounts was able to preserve the independently estimated means and variances of the cumulative rainfall amounts series.

APPENDIX A: DERIVATION OF THE CONDITIONAL PROBABILITY OF OCCURRENCE h_k

Let

$$h_k^{ij} = P\{Z_k = j | Z_0 = i\} \tag{A1}$$

Using combinatoric arguments, one can write:

$$h_k^{11} = a_1 \sum_{l=1}^{k-1} f_1(k-l) h_l^{11} + a_1 f_1(k) + (1-a_2) \sum_{l=1}^{k-1} f_2(k-l) h_l^{12} \tag{A2a}$$

$$h_k^{12} = (1-a_1) \sum_{l=1}^{k-1} f_1(k-l) h_l^{11} + (1-a_1) f_1(k) + a_2 \sum_{l=1}^{k-1} f_2(k-l) h_l^{12} \tag{A2b}$$

$$h_k^{21} = a_1 \sum_{l=1}^{k-1} f_1(k-l) h_l^{21} + (1-a_2) f_2(k) + (1-a_2) \sum_{l=1}^{k-1} f_2(k-l) h_l^{22} \tag{A2c}$$

$$h_k^{22} = (1-a_1) \sum_{l=1}^{k-1} f_1(k-l) h_l^{21} + a_2 f_2(k) + a_2 \sum_{l=1}^{k-1} f_2(k-l) h_l^{22} \tag{A2d}$$

where $f_1(\ )$ and $f_2(\ )$ have been defined in the text (equation (6)). Moreover,

$$h_k = e_1(h_k^{11} + h_k^{12}) + e_2(h_k^{21} + h_k^{22}) \tag{A3}$$

To simplify the algebra above, let us define

$$F_j(z) = f_j(1)z + f_j(2)z^2 + \cdots \qquad j = 1, 2 \tag{A4a}$$

$$H^{ij}(z) = h_1^{ij} z + h_2^{ij} z^2 + \cdots, \quad i = 1, 2 \quad j = 1, 2 \tag{A4b}$$

Then (A3) becomes

$$H(z) = e_1[H^{11}(z) + H^{12}(z)] + (1-e_1)[H^{21}(z) + H^{22}(z)] \tag{A5}$$

where, for example, $H^{11}(z)$ is given by

$$H^{11}(z) = a_1 F_1(z) H^{11}(z) + a_1 F_1(z) + (1-a_2) F_2(z) H^{12}(z) \tag{A6}$$

and similar expressions can be written for $H^{12}(z)$, $H^{21}(z)$, and $H^{22}(z)$. It is well known [cf. *Parzen*, 1960] that for a geometric distribution with parameter p,

$$F(z) = \frac{pz}{1-(1-p)z} \qquad j = 1, 2 \tag{A7}$$

After some algebra on (A5)–(A7), one can show that

$$H(z) = \frac{mz}{1-z} + \frac{Az}{1-Wz}$$

where m, A, and W have been defined in the text (equations (16), (19), and (20)). Therefore

$$h_k = m + AW^{k-1}$$

APPENDIX B: DERIVATION OF THE LIKELIHOOD FUNCTION

$$\begin{aligned} L(\theta | x_1, \cdots, x_n) &= P\{X_1 = x_1, \cdots, X_n = x_n\} \\ &= P\{X_1 = x_1, \cdots, X_n = x_n | \langle X_1 \rangle = 1\} e_1 \\ &\quad + P\{X_1 = x_1, \cdots, X_n = x_n | \langle X_1 \rangle = 2\}(1-e_1) \\ &= g_{1,n}{}^1 e_1 + g_{1,n}{}^2 (1-e_1) \end{aligned} \tag{B1}$$

where $e_1 = P\{\langle X_1 \rangle = 1\}$ is the unconditional probability of type 1 interval, and the conditional probabilities $g_{1,n}{}^1$ and $g_{1,n}{}^2$ are defined from the general formula

$$g_{i,j}{}^k = P\{X_i = x_i, \cdots, X_j = x_j | \langle X_i \rangle = k\} \tag{B2}$$

$$k = 1, 2 \qquad i = 1, \cdots, n-1 \qquad j = 2, \cdots, n$$

for $i = 1$ and $j = n$. It is easy to show that the conditional probabilities $g_{1,n}{}^1$ and $g_{1,n}{}^2$ can be written in a recursive form

$$g_{1,n}{}^1 = f_1(x_1)[g_{2,n}{}^1 a_1 + g_{2,n}{}^2 (1-a_1)] \tag{B3}$$

$$g_{1,n}{}^2 = f_2(x_1)[g_{2,n}{}^1 (1-a_2) + g_{2,n}{}^2 a_2] \tag{B4}$$

By further recursing expressions (B3) and (B4) and introducing matrix notation the likelihood function of the Markov renewal model takes the form of (24) given in the text.

Acknowledgments. We gratefully acknowledge the constructive comments of James A. Smith and anonymous reviewers on an earlier draft of this paper.

REFERENCES

Chang, T. J., M. L. Kavvas, and J. W. Delleur, Daily precipitation modeling by discrete autoregressive moving average processes, *Water Resour. Res.*, *20*(5), 565–580, 1984.

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

Cox, D. R., Some models for series of events, *Int. Stat. Inst.*, *40*, 737–746, 1963.

Cox, D. R., and V. Isham, *Point Processes*, Chapman and Hall, London, 1980.

Cox, D. R., and P. A. W. Lewis, *The Statistical Analysis of Series of Events*, Methuen, New York, 1978.

Cox, D. R., and H. D. Miller, *The Theory of Stochastic Processes*, Methuen, New York, 1965.

Foufoula-Georgiou, E., Discrete-time point process models for daily rainfall, Ph.D. dissertation, Univ. of Fla., Gainesville, 1985.

Foufoula-Georgiou, E., and P. Guttorp, Compatibility of continuous rainfall occurrence models with discrete rainfall observations, *Water Resour. Res.*, *22*(8), 1316–1322, 1986.

Foufoula-Georgiou, E., and D. P. Lettenmaier, Continuous-time versus discrete-time point process models for daily rainfall occurrences, *Water Resour. Res.*, *22*(4), 531–542, 1986.

Gabriel, K. R., and J. Neumann, A Markov chain model for daily rainfall occurrences at Tel Aviv, *Q. J. R. Meteorol. Soc.*, *88*, 90–95, 1962.

Galloy, E., A. LeBreton, and S. Martin, A model for weather cycles based on daily rainfall occurrence, in *Rhythmes in Biology and Other Fields of Application, Springer Lecture Notes in Biomathematics*, vol. 49, edited by M. Cosnard et al., pp. 303–318, 1981.

Gaver, D. G., and P. A. W. Lewis, First order autoregressive Gamma sequences and point processes, *Adv. Appl. Probl.*, *12*(3), 727–745, 1980.

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

Guttorp, P., On binary time series obtained from continuous time point processes describing rainfall, *Water Resour. Res.*, *22*(6), 897–904, 1986.

Jacobs, P. A., and P. A. W. Lewis, A discrete time series generated by mixtures, 1, Correlational and run properties, *J. R. Stat. Soc.*, *B40*(1), 94–105, 1978.

Kavvas, M. L., and J. W. Delleur, A stochastic cluster model for daily rainfall sequences, *Water Resour. Res.*, *17*(4), 1151–1160, 1981.

Keenan, D. M., *Time Series Analysis of Binary Data*, Ph.D. thesis, The Univ. of Chicago, Ill., 1980.

Lawrance, A. J., The mixed exponential solution to the first-order autoregressive model, *J. Appl. Probl.*, *17*, 546–552, 1980.

Leytham, K. M., Maximum likelihood estimates for the parameters of mixture distributions, *Water Resour. Res.*, *20*(7), 896–902, 1984.

Nelder, J. A., and R. Mead, A simplex method for function minimization, *Comput. J.*, *7*, 308–313, 1965.

Parzen, E., *Modern Probability Theory and its Applications*, John Wiley, New York, 1960.

Pyke, R., Markov renewal processes: Definitions and preliminary properties, *Ann. Math. Stat.*, *32*, 1231–1242, 1961*a*.

Pyke, R., Markov renewal processes with finitely many states, *Ann. Math. Stat.*, *32*, 1243–1259, 1961*b*.

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.

Smith, J. A., Statistical modeling of daily rainfall occurrences, *Water Resour. Res.*, in press, 1987.

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

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

Smith, W. L., Regenerative stochastic processes, *Proc. R. Soc. London, Ser. A*, *232*, 6–31, 1955.

Teugels, J. L., A bibliography on semi-Markov processes, *J. Comput. Appl. Math.*, *2*(2), 125–144, 1976.

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.

Waymire, E., and V. K. Gupta, The mathematical structure of rainfall representations, 2, A review of the theory of point processes, *Water Resour. Res.*, *17*(5), 1273–1286, 1981*a*.

Waymire, E., and V. K. Gupta, The mathematical structure of rainfall representations, 3, Some applications of the point process theory to rainfall processes, *Water Resour. Res.*, *17*(5), 1287–1294, 1981*b*.

Waymire, E., 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.

Woolhiser, D. A., and J. Roldan, Stochastic daily precipitation models, 2, A comparison of distributions of amounts, *Water Resour. Res.*, *18*(5), 1461–1468, 1982.

E. Foufoula-Georgiou, Department of Civil Engineering, Iowa State University, Ames, IA 50011.

D. P. Lettenmaier, Department of Civil Engineering, FX-10, University of Washington, Seattle, WA 98195.

(Received December 24, 1986;
revised February 12, 1987;
accepted February 13, 1987.)