WATER RESOURCES RESEARCH, VOL. 25, NO. 2, PAGES 295–302, FEBRUARY 1989

Probabilistic Representation of the Temporal Rainfall Process by a Modified Neyman-Scott Rectangular Pulses Model: Parameter Estimation and Validation

DARA ENTEKHABI, IGNACIO RODRIGUEZ-ITURBE,[1] AND PETER S. EAGLESON

Massachusetts Institute of Technology, Cambridge, Massachusetts

The capability of the Neyman-Scott clustered stochastic point process model of rainfall to preserve various observed statistics is considered. Randomization of the cell duration parameter from storm to storm is shown to considerably improve the wet-dry period, joint distribution, and extreme value statistics. A simple procedure for parameter estimation is introduced and applied.

Introduction

Stochastic processes consisting of point events occurring in time and having characteristics derived from sampling probability density functions are becoming well-established in hydrology. In particular, point process representations of rainfall occurrences have been used to drive various physically based models of basin storage and fluxes [*Eagleson*, 1978]. The sensitivity of these storages and fluxes to the temporal structure of the precipitation input has been well-established [e.g., *Eagleson*, 1978]. This sensitivity raises an important practical question. Do these models, fitted to cumulative precipitation data at one level of aggregation (say, daily), faithfully reproduce the statistics of the temporal structure at other important levels of aggregation (say, hourly)? The answer is that some models only perform well at the scale of aggregation for which they were constructed but other models will preserve the main statistical characteristics of the rainfall process over a relatively wide range of aggregation levels [*Rodriguez-Iturbe et al.*, 1984, 1988].

Models which are capable of an adequate representation of the rainfall process at a point through a range of temporal scales of aggregation are based on a clustered point process structure. Storm arrivals are assumed to follow a Poisson process. With each storm event is associated a random number of cells; natural candidates for the distribution of the number of cells are the Poisson distribution and the geometric distribution. Each cell is represented by a rectangular pulse of random intensity and duration. The positioning of the cells can be made in several different manners. Two natural ways of doing it lead to the Neyman-Scott process and the Bartlett-Lewis process [*Cox and Isham*, 1980]. In the Neyman-Scott process the positions of the cells are determined by a set of independent and identically distributed random variables representing the time intervals between the storm origin and the birth of the individual cells. In the Bartlett-Lewis process the intervals between successive cells are independent and identically distributed. Overlapping of cells is allowed both within cells of the same storm and across cells of different storms.

Rodriguez-Iturbe et al. [1987, 1988] (hereafter referred to as RCI (1987, 1988), studied in detail the characteristics of both the Neyman-Scott rectangular pulses process and the Bartlett-Lewis rectangular pulses process. The difference between the two is relatively subtle and it is very unlikely that empirical analysis of data can be used to choose between them. It was observed by RCI (1987) that both models are capable of preserving a number of the statistical characteristics of rainfall data at different levels of aggregation without changing the model parameters. Nevertheless, it was noticed that both models severely overestimate the probability of observed dry periods when those periods were above several hours. The implications of this for infiltration studies and for other hydrologic considerations such as rainfall runoff transformations is serious, since there can be a major difference in the runoff output when the period with no rainfall is varied.

The original versions of the cluster-based models considered rectangular cells whose stochastic description was invariant throughout the storm events. In other words, the duration of the cells, the intensity of the cells and the number of cells came from distribution functions whose parameters were the same for all storms. A modified version of the Bartlett-Lewis rectangular pulses model was developed by RCI (1988). Their modification allows for different structural characteristics among the different storms and is capable of representing a large variety of statistical characteristics of the rainfall process at different levels of aggregation including the probability of dry periods and other related properties. Similar changes to the Neyman-Scott rectangular pulses model are introduced in this paper.

A practical feature of the models described above is the efficiency of their parameter estimation procedures. Sensitivity analyses and empirical verification of such models require the repeated estimation of parameter sets based on different historical statistics estimated at different levels of aggregation. For this reason a simplified solution procedure to the set of nonlinear equations is introduced here with satisfactory convergence behavior.

Modified Neyman-Scott Rectangular Pulses Model

The Neyman-Scott rectangular pulses rainfall model introduced by RCI (1987) is a particular form of a clustered point process. There are storm origins that arive in a Poisson manner with parameter λ. With each storm are associated a

[1]Also at Instituto Internacional de Estudios Avanzados (IDEA), Caracas, Venezuela.

Paper number 88WR03502
0043-1397/89/88WR-03502$05.00

TABLE 1. Historical Estimates of Mean, Variance, and Lagged Autocorrelation of Cumulative Precipitation at Various Levels of Aggregation: Denver, Colorado, Period May 15 to June 16, 1949–1976

Level of Aggregation, hours	Mean, mm	Variance, mm^2	Lag-1 Autocorrelation	Lag-2 Autocorrelation	Lag-3 Autocorrelation
1	0.0885	0.4030	0.4800	0.3220	0.2679
6	0.5313	5.9702	0.3318	0.1285	0.0593
12	1.0626	16.8829	0.2301	0.0671	−0.0218
24	2.1252	41.6067	0.1571	−0.0250	−0.0473
48	4.2510	98.9673	0.0196	−0.0423	−0.0540

random c ($c \geq 1$) number of cells. The cells are each independently displaced from the storm origin according to an exponential probability density function with parameter β. Each cell is a rectangular pulse of random height (intensity) x and width (duration). The duration is exponentially distributed with parameter η. The probability density function for cell intensity need not be selected at this stage. There is the possibility of cell overlap both within and across storms.

In order to be able to equate the second-order properties of this process with those measured by standard rainfall gages, it is necessary to integrate the intensity process. The result Y of this aggregation over an arbitrary time period τ defines the rainfall depth process. As given by RCI (1987) the second-order properties of the aggregated original Neyman-Scott rectangular pulses process $Y_i(\tau)$, the cumulative rainfall over the time interval τ are

$$E[Y_i(\tau)] = \lambda\eta^{-1}E[c]E[x]\tau \tag{1}$$

$$\begin{aligned}\mathrm{Var}\,[Y_i(\tau)] = \lambda\eta^{-3}(\eta\tau - 1 + e^{-\eta\tau})\Big\{&2E[c]E[x^2] \\ &+ E[c^2 - c]E^2[x]\frac{\beta^2}{\beta^2 - \eta^2}\Big\} \\ &- \lambda(\beta\tau - 1 + e^{-\beta\tau})\frac{E[c^2 - c]E^2[x]}{\beta(\beta^2 - \eta^2)}\end{aligned} \tag{2}$$

$$\begin{aligned}\mathrm{Cov}\,[Y_i(\tau), Y_{i+k}(\tau)] &= \lambda\eta^{-3}(1 - e^{-\eta\tau})^2e^{-\eta(k-1)\tau} \\ &\cdot\left\{E[c]E[x^2] + \frac{1}{2}\frac{E[c^2 - c]E^2[x]\beta^2}{\beta^2 - \eta^2}\right\} - \lambda(1 - e^{-\beta\tau})^2 \\ &\cdot e^{-\beta(k-1)\tau}\frac{1}{2}\frac{E[c^2 - c]E^2[x]}{\beta(\beta^2 - \eta^2)} \qquad k \geq 1\end{aligned} \tag{3}$$

TABLE 2. Historical Estimates of Cumulative Precipitation Probabilities at Various Levels of Aggregation: Denver, Colorado, Period May 15 to June 16, 1949–1976

Level of Aggregation, hours	Probability		
	Depth = 0.0	Depth < 0.5 mm	Depth < 1.0 mm
1	0.9385	0.9569	0.9712
2	0.9169	0.9363	0.9553
4	0.8825	0.9068	0.9298
6	0.8550	0.8845	0.9117
12	0.7795	0.8206	0.8594
16	0.7315	0.7755	0.8233
24	0.6412	0.6979	0.7558
48	0.4769	0.5370	0.6181
96	0.2731	0.3287	0.4074
192	0.0833	0.1296	0.1574

One major shortcoming of the original Neyman-Scott rectangular pulses model outlined above is its inability to preserve the proportions of dry or wet periods especially on the scale of several hours to several days (RCI, 1987). Clearly, the one factor that controls the duration of precipitation and hence wet and dry runs is the inverse mean cell duration η. By introducing structural interstorm variability in η we seek greater physical realism and a more accurate preservation of historical fractions of wet periods. Elimination of the current overestimation of dry windows in the simulated series will improve the estimates of critical dependent hydrologic variables such as basin runoff and evaporation. This change leads to the modified Neyman-Scott model.

Instead of fixing η as a constant parameter which controls the distribution from which the duration of all cells arises, η is now a random variable which changes from storm to storm. Thus the duration of the cells from storm i are random quantities governed by an exponential distribution with parameter η_i. The probability density function for η is assumed to be a two-parameter gamma distribution with shape parameter α. All other assumptions remain the same as in the original Neyman-Scott rectangular pulses model and thus η is independent of the number of cells c and the cell intensities x. In the modified Bartlett-Lewis rectangular pulses model developed by RCI (1988) the same assumption is made for η but other parameters are also allowed to vary randomly from storm to storm. More specifically in the Bartlett Lewis case the parameter β of the exponential distribution controlling the cells interarrival time varies randomly from storm to storm but the dimensionless parameter β/η remains fixed. Also, in the Bartlett-Lewis case by assuming the number of cells per storm is geometrically distributed the process of cell origins in a storm terminates after a time that is exponentially distributed with rate γ, where $E[c] = 1 + \beta/\gamma$. In the modified Bartlett-Lewis rectangular pulses model γ also varies randomly from storm to storm but γ/η remains fixed. The Bartlett-Lewis rectangular pulses scheme is more tractable from a mathematical point of view than the Neyman-Scott scheme and thus a wider spectrum of analytical results are derived in the work by RCI (1988).

A very interesting feature of the cluster-based rectangular pulses models with random η is that for values of $\alpha < 2$ the processes become asymptotically second-order self-similar (RCI, 1988). This means that for large periods of aggregation the correlation structure converges to the correlation structure of the aggregated fractional noise process. The validity or lack of validity of this kind of behavior in temporal rainfall data aggregated for intervals between 1 hour and 1 month is

TABLE 3. Definition of Parameter Sets a, b, and c for the Original and Modified Neyman-Scott Rectangular Pulses Models

	Mean	Variance	Lag-1 Autocorrelation
Original Neyman-Scott Rectangular Pulses Model			
Parameter set a	1 hour	1 and 6 hours	1 and 6 hours
Parameter set b	1 hour	1 and 12 hours	1 and 12 hours
Parameter set c	1 hour	1 and 24 hours	1 and 24 hours
Modified Neyman-Scott Rectangular Pulses Model			
Parameter set a	1 hour	1 and 12 hours	1, 6, and 12 hours
Parameter set b	1 hour	1 and 24 hours	1, 6, and 24 hours
Parameter set c	1 hour	1 hour	1, 6, 12, and 24 hours

studied in detail by *Jacobs et al.* [1987]. Thus the models are structurally capable of representing either a fractal or a nonfractal type of process and this choice is not an a priori decision but rests upon the data from which the model parameters are estimated.

Using standarad properties of conditional expectation

$$E(Y) = E_\eta[E(Y\mid\eta)] \tag{4}$$

$$\mathrm{Var}\,(Y) = \mathrm{Var}\,[E(Y\mid\eta)] + E[\mathrm{Var}\,(Y\mid\eta)] \tag{5}$$

$$\mathrm{Cov}\,[Y_1, Y_2] = \mathrm{Cov}\,[E(Y_1\mid\eta), E(Y_2\mid\eta)] + E[\mathrm{Cov}\,(Y_1, Y_2\mid\eta)] \tag{6}$$

and assuming $\eta^2 \gg \beta^2$, the moments of the cumulative rainfall process are derived as for (1)–(3):

$$E[Y_i(\tau)] = E[x]E[c]\lambda\tau I(1, 0) \tag{7}$$

$$\begin{aligned}\mathrm{Var}\,[Y_i(\tau)] = {}& [E[x]E[c]\lambda\tau I(1, 0)]^2 \\ & + \{2C_1\tau + C_2\beta^{-3}(\beta\tau + e^{-\beta\tau} - 1) \\ & + (E[c]E[x]\lambda\tau)^2\}I(2, 0) \\ & - 2C_1 I(3, 0) - C_2\tau I(4, 0) + C_2 I(5, 0) \\ & + 2C_1 I(3, \tau) - C_2 I(5, \tau)\end{aligned} \tag{8}$$

$$\begin{aligned}\mathrm{Cov}\,[Y_i(\tau), Y_{i+k}(\tau)] = {}& C_1 I(3, k\tau - \tau) - 2C_1 I(3, k\tau) \\ & + C_1 I(3, k\tau + \tau) - \frac{C_2}{2} I(5, k\tau - \tau) \\ & + C_2 I(5, k\tau) - \frac{C_2}{2} I(5, k\tau + \tau) \\ & + \frac{C_2}{2}\beta^{-3}(1 - e^{-\beta\tau})^2 e^{-\beta(k-1)\tau} I(2,0) \\ & + (E[c]E[x]\lambda\tau)^2[I(2, 0) - I^2(1, 0)]\end{aligned} \tag{9}$$

where

$$C_1 = \lambda E[c]E[x^2] \tag{10}$$

$$C_2 = \lambda E[c^2 - c]E^2[x]\beta^2 \tag{11}$$

$$I(x, y) = E[\eta^{-x}e^{-\eta y}] = \frac{\Gamma(\alpha - x)}{\Gamma(\alpha)}\,\theta^{\alpha}(\theta + y)^{x-\alpha} \qquad x > 0,\ y \geq 0 \tag{12}$$

To evaluate the expectations $E[x]$, $E[x^2]$, $E[c]$, and $E[c^2 - c]$ we must assume distributions for x and c. For x exponentially distributed,

$$E[x] = \mu_x \tag{13}$$

$$E[x^2] = 2\mu_x^2 \tag{14}$$

For $c \geq 1$, the distribution may be either geometric or Poisson, in which case

$$E[c] = \mu_c \tag{15}$$

$$E[c^2 - c] = \delta\mu_c^2 + 2\varepsilon\mu_c \tag{16}$$

where

$$\begin{aligned}\delta &= +1 \quad c\text{: Poisson} \\ \delta &= +2 \quad c\text{: geometric}\end{aligned} \tag{17}$$

$$\begin{aligned}\varepsilon &= +1 \quad c\text{: Poisson} \\ \varepsilon &= -1 \quad c\text{: geometric}\end{aligned} \tag{18}$$

It can be shown by simulation that the distribution of the number of cells per storm has no general bias effect on the covariance and remaining functions [*Rodriguez-Iturbe et al.*, 1987]. Hereafter and in the simulations x is regarded as exponentially distributed and c is assumed to belong to a Poisson probability mass function. The random variable c is independently and identically distributed (iid) among the storms. The cell intensity and displacement from storm origins are iid among the cells of all storms. The cell duration is iid according to an exponential distribution with parameter η within every storm and η is gamma distributed between storms.

TABLE 4. Summary of Estimated Parameters

	Original Neyman-Scott Process				
	λ, hour^{-1}	β, hour^{-1}	η, hour^{-1}	$E[x]$, mm hour^{-1}	$E[c]$
Set a	0.005874	0.144365	2.072598	3.011935	10.36733
Set b	0.005070	0.179066	2.731938	3.052653	15.62309
Set c	0.005084	0.127560	2.102091	3.024823	12.09678

	Modified Neyman-Scott Process					
	λ, hour^{-1}	β, hour^{-1}	θ, hour	α	$E[x]$, mm hour^{-1}	$E[c]$
Set a	0.0104	0.1368	1.2212	4.0288	3.2441	6.2755
Set b	0.0107	0.1493	1.1304	3.9779	3.1398	6.9829
Set c	0.0109	0.1398	1.0625	3.7785	3.1833	6.5691

TABLE 5*a*. Set a Statistics for the Original Neyman-Scott Rectangular Pulses Model

Level	Mean	Variance	Corr(1)	Corr(2)	Corr(3)
1	0.08850	0.40300	0.48000	0.24402	0.18965
6	0.53100	5.97027	0.33179	0.13093	0.05506
12	0.16200	15.90236	0.24355	0.04171	0.00738
24	2.12401	39.55064	0.13444	0.00411	0.00013

TABLE 5*c*. Set c Statistics for the Original Neyman-Scott Rectangular Pulses Model

Level	Mean	Variance	Corr(1)	Corr(2)	Corr(3)
1	0.08850	0.40300	0.47999	0.28255	0.22849
6	0.53100	6.25134	0.35035	0.11577	0.03954
12	1.06200	16.88304	0.23010	0.02634	0.00307
24	2.12400	41.53556	0.11619	0.00156	0.00002

PARAMETER ESTIMATION

The parameters of the probability distribution functions governing the clustered Poisson-based models are estimated by equating their analytical moments with the moments estimated from historical precipitation time series. Equations (1)–(3) and (7)–(9) represent the mean, variance, and lagged covariance of processes defined by the original and modified Neyman-Scott rectangular pulse models, respectively. They are integrated equations over nonoverlapping time intervals of length τ in order to correspond to the manner in which rainfall is measured. Tables 1 and 2 contain the estimated statistical properties of rainfall in Denver, Colorado, for the seasonal period between May 15 and June 16. This period was concluded to be temporally homogeneous in regard to the rainfall characteristics by *Cordova and Bras* [1981]. The tables are based on 27 years of observations occurring between 1949 and 1976.

The objective is to estimate the $(n \times 1)$ parameter vector ν. Let $\hat{\mathbf{f}} = (\hat{f}_1, \hat{f}_2, \cdots, \hat{f}_n)$ be some estimated moments at different levels of aggregation from the historical data set (Table 1). The function $\mathbf{f}(\nu) = [f_1(\nu), f_2(\nu), \cdots, f_n(\nu)]$ contains the model expressions for the same moments based on the parameter vector ν. It is required that

$$\mathbf{f}(\nu) - \hat{\mathbf{f}} = \mathbf{0} \tag{19}$$

where $\mathbf{0}$ is the null vector. Equation (19) defines a system of n nonlinear simultaneous equations in n unknowns.

Multivariate Newton-Raphson solution of (19) requires the repeated inversion of Jacobian matrices. This approach and its modified versions that bypass the inversion requirement at every iteration are highly sensitive to the initial estimates of ν. They rarely converge and thus present an unreliable approach to the systematic estimation of parameters.

The alternative is to estimate ν by minimizing the sum of squared differences in (19). The components must be normalized so that the minimization is not biased toward preferentially selecting the larger magnitude components. For this, we define the diagonal matrices $\hat{\mathbf{F}} = \text{diag}(\hat{\mathbf{f}})$ and $\mathbf{F}(\nu) = \text{diag}[\mathbf{f}(\nu)]$. The minimization now becomes

$$\min_{\nu} z = \text{trace}\,\{(\mathbf{I} - \mathbf{F}(\nu)\hat{\mathbf{F}}^{-1})^2\} \tag{20}$$

Equation (20) poses a nonlinear unconstrained optimization problem. Solutions may be obtained using widely available algorithms. We have found (20) to be a reliable and rapid approach to parameter estimation of the Neyman-Scott rectangular pulses models and similar stochastic models.

The method of (20) is applied to the original Neyman-Scott rectangular pulses model for three different sets (a, b, and c) of five equations representing combinations of hourly mean, variance and lag-one autocorrelation and either the 6-, 12-, or 24-hourly variance and lag-one autocorrelation (see Table 3). The five equations in each case are simultaneously satisfied for their historical values and three sets of the five model parameters are defined. Similar estimation of ν is performed for the modified Neyman-Scott rectangular pulses model (equations (7)–(9)). There are six parameters in this case one of which is the probability distribution shape parameter α. If one is willing to guess the shape of the distribution, α may be set a priori at a constant value whereupon the modified and original Neyman-Scott rectangular pulses models will have the same number of parameters. In this experiment, the shape parameter is allowed to vary freely. The resulting parameter sets for both models are summarized in Table 4.

With these parameter sets, the moments are computed for both models according to (1)–(3) and (7)–(9), respectively, as a check on the goodness of fit. These values are given in Tables 5*a*, 5*b*, and 5*c* for the original Neyman-Scott model and in Tables 6*a*, 6*b*, and 6*c* for the modified version. Their agreement with the historical statistics given in Table 1 is good, indicating close satisfaction of (19). Two further observations are in order: (1) the values of the probability distribution parameters remain nearly the same regardless of which of the three constraint vectors $\hat{\mathbf{f}}$ are considered and (2) the model estimates of mean, variance, and lagged autocorrelation at aggregation levels *not* included in $\hat{\mathbf{f}}$ and $\mathbf{f}(\nu)$ are well-represented by the models. Therefore we can say that the rainfall process characterized by the statistics of Table 1 has a temporal structure consistent with a Neyman-Scott point process and that the suggested parameter estimation scheme is robust. Once the parameters are estimated, the statistics for every aggregation level in the stochastic process are constrained to the value derived from the analytical relations in (1)–(3) and (7)–(9).

Note that α is always more than 2, meaning that the model is not self-similar. The aggregated data from 1 to 24 hours do

TABLE 5*b*. Set b Statistics for the Original Neyman-Scott Rectangular Pulses Model

Level	Mean	Variance	Corr(1)	Corr(2)	Corr(3)
1	0.08850	0.40300	0.48000	0.24919	0.19817
6	0.53100	6.03365	0.35465	0.15575	0.07245
12	1.06200	16.34698	0.27262	0.05741	0.01242
24	2.12400	41.60698	0.15710	0.00722	0.00034

TABLE 6*a*. Set a Statistics for the Modified Neyman-Scott Rectangular Pulses Model

Level	Mean	Variance	Corr(1)	Corr(2)	Corr(3)
1	0.0846	0.3936	0.4776	0.2549	0.1887
6	0.5074	6.2653	0.2992	0.1246	0.0640
12	1.0148	17.3089	0.2216	0.0602	0.0327
24	2.0296	46.4089	0.1398	0.0433	0.0411

TABLE 6*b*. Set b Statistics for the Modified Neyman-Scott Rectangular Pulses Model

Level	Mean	Variance	Corr(1)	Corr(2)	Corr(3)
1	0.0886	0.4030	0.4800	0.2633	0.1972
6	0.5317	6.5179	0.3022	0.1199	0.0590
12	1.0634	18.1055	0.2163	0.0554	0.0319
24	2.1269	48.5688	0.1339	0.0441	0.0434

not indicate self-similarity under the criteria developed by *Cox* [1984] and applied to rainfall time series by [B. L. Jacobs, personal communication, 1987].

DRY PERIODS AND RELATED STATISTICAL PROPERTIES

Following the parameter estimation approximately 10^5 values of hourly precipitation depths are simulated for each of the three different parameter sets and for each model.

As previously noted, the original Neyman-Scott rectangular pulses model grossly overestimates the proportion of dry periods to the length of the series, especially for periods of aggregation ranging from several hours to a few days (RCI, 1987). This can be seen in the upper portion of Table 7 where the probability of dry 1-, 6-, 24-, and 48-hour periods are 0.9385, 0.8550, 0.6412, and 0.4769 in the historical data set, while the original model predicts values averaging 0.9498, 0.8793, 0.7788, and 0.6842, respectively.

The improvement introduced by the modification is tabulated in the lower portion of Table 7 and is strikingly displayed graphically in Figure 1. In this figure the probability estimates from the historical data set are plotted as separate points while the values from the simulations are plotted as the dotted (original model) and solid (modified model) continuous curves. Even up to 8-day aggregation periods the agreement between the historical probabilities and the modified process probabilities are good. In addition to preserving the mean, variance and lagged autocorrelation, the modified Neyman-Scott rectangular pulses model preserves the dry-wet time structure of point observations of rainfall. These characteristics are particularly important in the hydrologic application of such models where estimates of runoff storage, and evaporation of a basin are of interest.

It is interesting to notice that in the original and modified Bartlett-Lewis rectangular pulses models the probability of dry periods or what is equivalent, the probability mass at the origin of the aggregated process, is analytically known and thus the expression can be used in the parameter estimation procedure (RCI, 1987). In the Neyman-Scott schemes the

TABLE 6*c*. Set c Statistics for the Modified Neyman-Scott Rectangular Pulses Model

Level	Mean	Variance	Corr(1)	Corr(2)	Corr(3)
1	0.0873	0.3976	0.4791	0.2605	0.1942
6	0.5241	6.3994	0.3017	0.1231	0.0618
12	1.0481	17.7588	0.2197	0.0576	0.0318
24	2.0963	47.7142	0.1365	0.0435	0.0433

TABLE 7. Probability of No Events at Various Levels of Aggregation for the Simulated Data

Level of Aggregation, hours	Observed (Denver)	Simulated With Parameter Set a	b	c
Original Neyman-Scott Rectangular Pulses Model				
1	0.9385	0.9504	0.9496	0.9493
2	0.9169	0.9258	0.9290	0.9251
4	0.8825	0.8960	0.9055	0.8951
6	0.8550	0.8756	0.8882	0.8742
12	0.7795	0.8329	0.8540	0.8338
16	0.7315	0.8077	0.8343	0.8127
24	0.6412	0.7697	0.7955	0.7713
48	0.4769	0.6675	0.7050	0.6800
96	0.2731	0.5040	0.5480	0.5290
192	0.0833	0.2920	0.3380	0.3220
Modified Neyman-Scott Rectangular Pulses Model				
1	0.9385	0.9406	0.9360	0.9368
2	0.9169	0.9042	0.8996	0.8985
4	0.8825	0.8545	0.8500	0.8471
6	0.8550	0.8191	0.8163	0.8106
12	0.7795	0.7494	0.7437	0.7360
16	0.7315	0.7097	0.7090	0.7003
24	0.6412	0.6480	0.6455	0.6305
48	0.4769	0.4990	0.4980	0.4835
96	0.2731	0.2990	0.2980	0.2940
192	0.0833	0.0780	0.0980	0.1000

analytical expression is not available and without making use of this information in the estimation of parameters, the modified model, indeed, preserves well the dry period characteristics of the historical data at different levels of aggregation.

For further assessment of fit several joint distribution characteristics are evaluated both for the historical data and via simulation for the original and modified Neyman-Scott rectangular pulses models. The results are shown in Table 8 where the fixed η model refers to the set a of parameters specified in Table 4.

The results are similar to those found by RCI (1987) for the random η Bartlett-Lewis rectangular pulses model. The

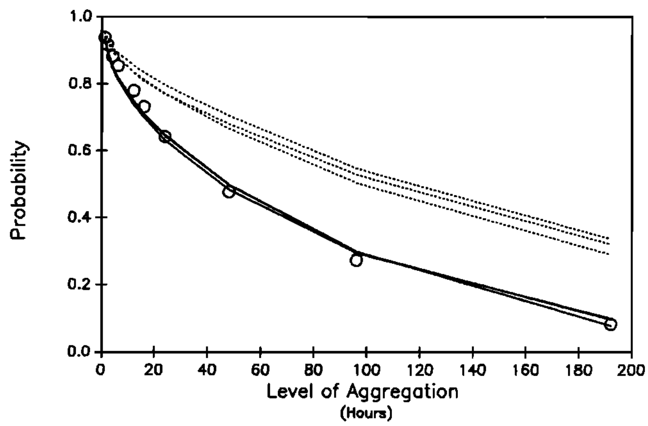

Fig. 1. Probability of zero depth from simulations with the original Neyman-Scott rectangular pulses model parameter sets a, b, and c (dashed curves) and with the modified Neyman-Scott rectangular pulses model parameter sets a, b, and c (solid curves). The historical values for Denver, Colorado (May 16, to June 16, 1949–1976) are represented by open circles.

TABLE 8. Models and Historical Comparison of Statistics That Characterize the Temporal Structure of Rainfall

	1 hour	6 hours	12 hours	24 hours
	$Prob\ (Y_t = 0 \mid Y_{t-1} = 0)$			
Historical	0.976	0.901	0.820	0.712
Original N-S	0.974	0.950	0.923	0.863
Modified N-S	0.961	0.913	0.866	0.773
	$E(Y_t \mid Y_t > 0)$			
Historical	1.454	3.692	4.844	5.943
Original N-S	1.874	4.477	6.667	9.678
Modified N-S	1.510	2.975	4.295	6.115
	$Var^{1/2}\ (Y_t \mid Y_t > 0)$			
Historical	2.158	5.426	7.684	9.649
Original N-S	2.499	5.419	7.298	9.587
Modified N-S	2.399	5.268	7.252	9.628
	$E(Y_t \mid Y_t > 0,\ Y_{t-1} = 0)$			
Historical	1.240	2.910	2.811	4.461
Original N-S	1.743	4.909	8.330	12.537
Modified N-S	1.309	3.074	4.831	7.127
	$Var^{1/2}\ (Y_t \mid Y_t > 0,\ Y_{t-1} = 0)$			
Historical	2.318	5.061	6.894	7.896
Original N-S	2.381	5.759	8.183	9.995
Modified N-S	2.106	5.038	6.806	10.281
	$E(Y_t \mid Y_t > 0,\ Y_{t+1} = 0)$			
Historical	0.845	2.271	3.002	4.131
Original N-S	1.323	2.758	4.368	8.236
Modified N-S	1.047	2.100	3.246	5.205
	$Var^{1/2}\ (Y_t \mid Y_t > 0,\ Y_{t+1} = 0)$			
Historical	1.257	3.950	4.882	6.799
Original N-S	1.798	4.244	6.302	9.513
Modified N-S	1.665	3.292	4.620	8.719
	$E(Y_t \mid Y_t > 0,\ Y_{t-1} > 0)$			
Historical	1.576	4.445	6.115	6.983
Original N-S	2.003	4.243	5.618	7.267
Modified N-S	1.826	2.906	3.926	5.370
	$Var^{1/2}\ (Y_t \mid Y_t > 0,\ Y_{t-1} > 0)$			
Historical	2.264	5.887	8.697	10.602
Original N-S	2.605	5.213	6.476	8.532
Modified N-S	2.772	5.412	7.517	9.057
	$Corr\ [Y_t, Y_{t+1} \mid Y_t > 0,\ Y_{t+1} > 0]$			
Historical	0.338	0.508	0.317	0.100
Original N-S	0.206	0.110	−0.0018	−0.137
Modified N-S	0.492	0.303	0.235	0.124

N-S, Neyman-Scott.

modified Neyman-Scott rectangular pulses model, indeed, preserves very well the different joint distribution characteristics that were analyzed for a relatively wide range of levels of aggregation. It shows in most joint features a striking improvement over the original fixed η version of the model.

It is important to point out that there are climates where nonstationarities are present at time scales lower than 1 day [*Maddox*, 1980]. This is frequently the case where rainfall comes mainly from convective showers whose air lifting due to heating plays a dominant role. In these cases a high percentage of the precipitation is concentrated in the afternoon and evening hours and thus one may find a higher autocorrelation in the daily rainfall totals than, say, the 12-hour accumulated rainfall. The models in their present structure can not properly represent this type of behavior, and the insertion of a cyclic structure in the rate of storm arrivals λ is a logical approach to represent this kind of physical situation.

From the above simulations an extreme value analysis was carried out for the original and the random η Neyman-Scott rectangular pulses models. The results are shown in Figures 2*a* and 2*b* where they are compared with the extremes of the historical data. In the historical case the largest hour depth for each of the 32 days in each year is picked and assigned a probability according to its rank within the sorted sample. A similar procedure is followed for the simulated series but with a much larger sample. The experiments of RCI (1987) show that the original version of the model adequately reproduce the extreme value behavior of the historic data. The results obtained in this paper show an adequate behavior in this aspect for the random η version of the model as well along with an improved representation of the rare event tail.

SUMMARY AND CONCLUSIONS

A modified Neyman-Scott rectangular pulses point process is proposed wherein the cell durations are iid according to an exponential probability distribution with parameter η within every storm. The parameter η is iid between storms according to a gamma probability density function.

A method of moments parameter estimation is followed where a set of highly nonlinear equations is solved by a convenient nonlinear unconstrained optimization in parameter vector ν.

Following parameter estimation for both the original and modified rectangular pulses Neyman-Scott models using various mixes of historical statistics to be preserved in the method of moments estimations, long series are synthetically generated. Whereas the original model severely overestimates periods of no rainfall, the modified model preserves the historical probabilities of dry periods up to 8 days in duration and also better preserves the statistical characteristics of the joint distribution for two successive periods.

The internal structure of rainfall events lies hidden in the aggregated statistics but is critical in the establishment of infiltration, runoff, moisture, and energy balance, and other hydrologic and climatologic variates. Stochastic point process models such as the random η Neyman-Scott and Bartlett-Lewis rectangular pulses models may hence become valuable tools in the deciphering of constituent components from integrated observations. More research, however, is needed in determining the comparative robustness of the parameter estimates when only large aggregation periods (e.g., 24 and 12 hours) are used in the estimation process. Nevertheless, it is clear that these models accomplish a consistent representation of the rainfall process at a continuum of scales ranging from less than an hour to more than a few days. Moreover, our research in progress shows that when these models are applied to different months of the year and to regions with and without orographic influences, the estimated model parameters reflect the distinguishing climatologic features of the season and the location.

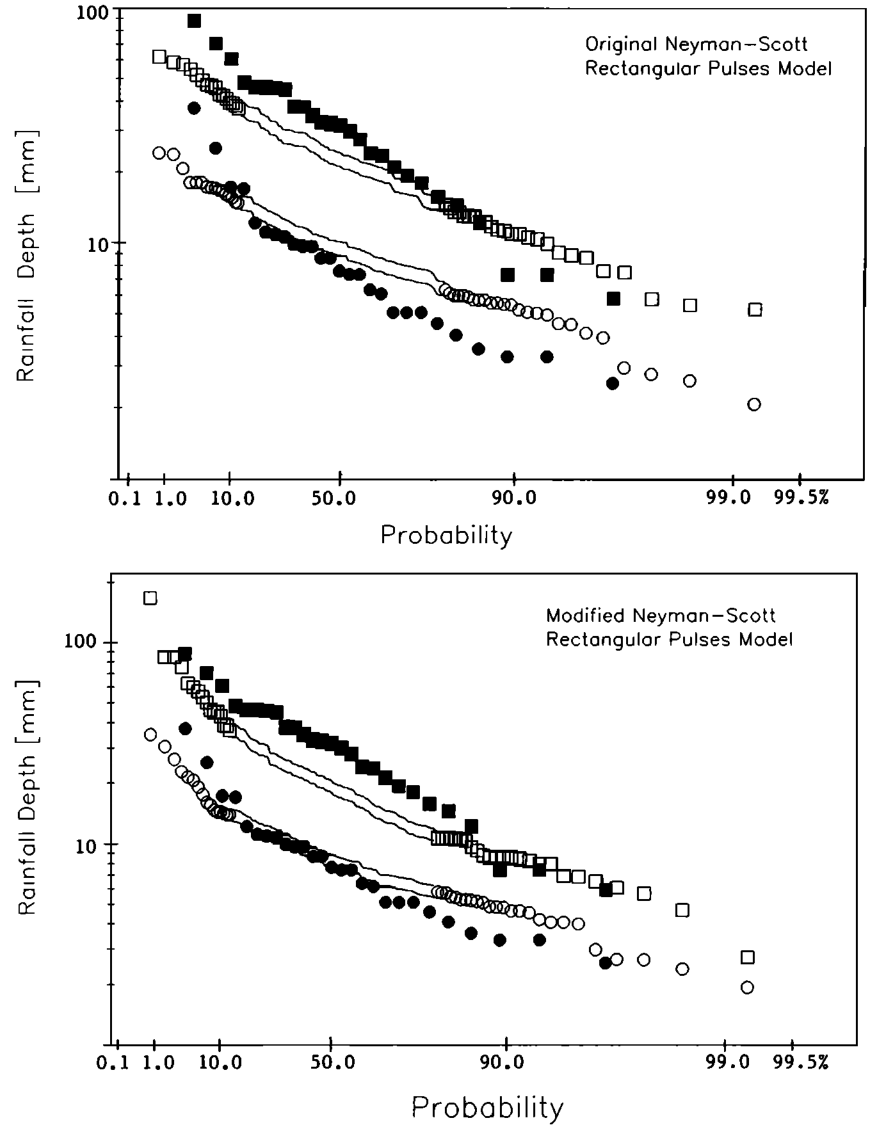

Fig. 2. Extreme value analysis for the 1-hour (circles) and 24-hour (squares) levels of aggregation. The open symbols represent results from simulation with parameter set a. Closed symbols are the Denver, Colorado (May 15, to June 16, 1949–1976), historical values.

NOTATION

- c number of cells per storm.
- $E[\ ]$ expectation operator.
- $\mathbf{f}(\nu)$ vector of analytical expressions for statistical moments as functions of parameter vector ν.
- $\hat{\mathbf{f}}$ vector of observed statistical moments.
- $F(\nu)$ matrix with $\mathbf{f}(\nu)$ as diagonal.
- $\hat{\mathbf{F}}$ matrix with $\hat{\mathbf{f}}$ as diagonal.
- $I(\ )$ analytical function defined by (12), hours.
- k time lag, hour.
- $\mathrm{Var}(\ )$ variance of ( ), mm^2.
- x cell intensity, $mm/hour^{-1}$.
- $Y(\tau)$ cumulative rainfall process aggregated over τ hours, mm.
- α shape parameter of cell duration gamma probability density function (pdf).
- β exponential pdf parameter of cell interarrival, $hour^{-1}$.
- δ constant depending on pdf of c, (17).
- ε constant depending on pdf of c, (18).
- η exponential pdf parameter of cell duration, $hour^{-1}$.
- θ parameter of cell duration gamma pdf.
- λ parameter of storm center interarrival exponential pdf, $hour^{-1}$.
- μ_c expectation of c.
- μ_x expectation of x, $mm/hour^{-1}$.
- ν parameter vector.
- τ period of aggregation, hours.

Acknowledgment. This material is based upon work supported in part by the National Science Foundation under grant ATM-8420781.

REFERENCES

Cordova, J. R., and R. L. Bras, Physically based probabilistic models of infiltration, soil moisture and actual evapotranspiration, *Water Resour. Res.*, *17*(1), 93–106, 1981.

Cox, D. R., Long-range dependence: A review, in *Statistics: An Appraisal*, edited by H. A. David and H. T. David, Iowa State University Press, Ames, 1984.

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

Eagleson, P. S., Climate, soil and vegetation, 7, A derived distribution of annual water yield, *Water Resour. Res.*, *14*(5), 765–776, 1978.

Jacobs, B. L., P. S. Eagleson, and I. Rodriguez-Iturbe, Stochastic modeling of precipitation events in space and time: Parameter estimation and scales of fluctuations, *R. M. Parsons Lab. Rep. 314*, Dep. of Civ. Eng., Mass. Inst. of Technol., Cambridge, 1987.

Maddox, R. A., Mesoscale convective complexes, *Bull. Am. Meteorol. Soc.*, *61*, 1374–1387, 1980.

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 V. Isham, Some models for rainfall based on stochastic point processes, *Proc. R. Soc. London, Ser. A*, *410*, 269–288, 1987.

Rodriguez-Iturbe, I., D. R. Cox, and V. Isham, A point process model for rainfall: Further developments, *Proc. R. Soc. London, Ser. A*, *417*, 283–298, 1988.

Rodriguez-Iturbe, J., B. Febres de Power, and J. B. Valdés, Rectangular pulses point process models for rainfall: Analysis of empirical data, *Geophys. Res.*, *92*(D8), 9645–9656, 1987.

P. S. Eagleson, D. Entekhabi, and I. Rodriguez-Iturbe, Ralph M. Parsons Laboratory, Department of Civil Engineering, Building 48-335, Massachusetts Institute of Technology, Cambridge, MA 02139.

(Received February 16, 1988;
revised August 12, 1988;
accepted August 19, 1988.)