WATER RESOURCES RESEARCH, VOL. 30, NO. 12, PAGES 3423–3435, DECEMBER 1994

Aggregation-disaggregation properties of a stochastic rainfall model

Zhiquan Bo and Shafiqul Islam

Department of Civil and Environmental Engineering, University of Cincinnati, Cincinnati, Ohio

E. A. B. Eltahir

Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

Abstract. A statistical approach based on the modified Bartlett-Lewis rectangular pulses model is presented to disaggregate rainfall statistics from daily data. Six model parameters are estimated from 24- and 48-hour accumulated rainfall data. Based on these estimated parameters, in addition to reproducing 24- and 48-hour statistics, the model is shown to infer 1-, 2-, 6-, and 12-hour historical statistics satisfactorily. An upper limit for disaggregation scale (about 2 days) for this model has been identified. This characteristic behavior of the model is related to the power law dependence of the power spectrum for timescales smaller than 2 days. A detailed comparison between observed and modeled statistics of rainfall data is presented for two rain gages, one from central Italy and the other from the midwestern United States.

1. Introduction

Precipitation data is widely collected by rain gages for nonoverlapping intervals such as seconds, minutes, hours, days, etc. However, most of the precipitation data are archived on a daily timescale. For hydrologic applications, for example, in modeling rainfall-runoff transformations, using time-dependent infiltration models, and simulating wetland dynamics, we need rainfall data at a much finer timescale. Recently, methods of disaggregating daily rainfall into shorter time periods have been proposed [e.g., *Srikanthan and McMahon*, 1985; *Hershenhorn and Woolhiser*, 1987; *Econopouly et al.*, 1990]. These models attempt to disaggregate daily rainfall into sequences of showers. However, they require dozens of parameters to disaggregate daily rainfall into individual storms. In this paper we present a different approach to disaggregate daily rainfall into a sequence of storms. Instead of trying to reproduce the specific rainfall events, we attempt to capture the statistics of finer-scale (e.g., 1 and 2 hours) rainfall from the observed daily rainfall statistics. Once the parameters are estimated, we can simulate the sequence of rainfall events at any desired timescale (e.g., from seconds to days). Our approach is based on a point process model developed and modified by *Rodriguez-Iturbe et al.* [1987 and 1988]. Using finer-resolution rainfall data, this model has been shown to capture temporal and spatial structure of rainfall [*Islam et al.*, 1990; *Onof and Wheater*, 1993]. In this paper we will investigate the disaggregation properties of this model. Here, disaggregation is used in reference to using daily accumulated rainfall statistics to infer finer-scale statistics, while aggregation refers to using finer-scale (1 and 2 hours) statistics to infer daily rainfall statistics. Specifically, we want to address the following two questions in this paper:

Paper number 94WR02026.
0043-1397/94/94WR-02026$05.00

1. Can we use readily available daily rainfall data to infer the finer-timescale (e.g., hourly) statistics?
2. Can we identify any specific structure of the modified Bartlett-Lewis rectangular pulses model (MBRPM) that explains the model's ability (or inability) to disaggregate rainfall statistics at various temporal scales?

A brief description of the model structure and the expressions for second-order moments is presented in section 2. Two data sets used for this study are described in section 3. Parameter estimation methodology and measures of goodness of fit are described in section 4. In section 5 the power spectrum of the rainfall model is derived and used to explain the results of disaggregation procedure. Concluding remarks are given in section 6.

2. Model Description

For this study, we use a point process based stochastic model of rainfall [*Rodriguez-Iturbe et al.*, 1987, 1988; *Islam et al.*, 1990]. This model, commonly known as the Bartlett-Lewis rectangular pulses rainfall model, is based on a cluster-based Poisson arrival process of storm origins with a rate λ. Each storm is characterized by a random number of cells C ($C \geq 1$), and each storm origin is followed by a Poisson arrival, at rate β, of cell origins. The intervals between successive cells are independent and identically distributed random variables. The candidate for this distribution is an exponential distribution with rate γ until the process of cell origin ends. Each cell itself is a rectangular pulse of random height (intensity) and width (duration). The cell duration has an exponential distribution with parameter η. For mathematical convenience, two dimensionless parameters $\kappa = \beta/\eta$ and $\phi = \gamma/\eta$ are introduced. The number of cells per storm, C, thus has a geometric distribution of mean $\mu_C = 1 + \kappa/\phi$. The original Bartlett-Lewis model is extended by *Rodriguez-Iturbe et al.* [1988] to capture the observed higher degree of correlation between the duration of cells within a single storm and to better reproduce the probability of zero rainfall. In the extended model, known as the modified Bartlett-Lewis model, the cell duration parameter η is taken to be a random variable that changes from storm to storm. The probability density function for η is assumed to be a two-parameter (α and ν) gamma distribution with shape parameter α. Each cell depth is a random constant exponentially distributed with mean $E[x]$. All the random variables defining the process are assumed to be mutually independent. Following *Rodriguez-Iturbe et al.* [1987, 1988], the second-order properties of the accumulated process over the time interval T, $Y_t^{(T)}$ for the modified Bartlett-Lewis rainfall model are summarized below:

$$E[Y_t^{(T)}] = \lambda E[x]\mu_C \frac{\nu}{\alpha - 1} T \tag{1}$$

$$\begin{aligned}\mathrm{Var}\,[Y_t^{(T)}] = \frac{2\nu^{2-\alpha}T}{\alpha-2}\left(k_1 - \frac{k_2}{\phi}\right) - \frac{2\nu^{3-\alpha}}{(\alpha-2)(\alpha-3)} \\ \cdot\left(k_1 - \frac{k_2}{\phi^2}\right) + \frac{2}{(\alpha-2)(\alpha-3)} \\ \cdot\left[k_1(T+\nu)^{3-\alpha} - \frac{k_2}{\phi^2}(\phi T + \nu)^{3-\alpha}\right]\end{aligned} \tag{2}$$

$$\begin{aligned}\mathrm{Cov}\,[Y_t^{(T)}, Y_{t+s}^{(T)}] &= \gamma_Y^T(s) \\ &= \frac{k_1}{(\alpha-2)(\alpha-3)}\{[T(s-1)+\nu]^{3-\alpha} \\ &\quad + [T(s+1)+\nu]^{3-\alpha} - 2(Ts+\nu)^{3-\alpha}\} \\ &\quad + \frac{k_2}{\phi^2(\alpha-2)(\alpha-3)}\{2(\phi Ts+\nu)^{3-\alpha} \\ &\quad - [\phi T(s-1)+\nu]^{3-\alpha} \\ &\quad - [\phi T(s+1)+\nu]^{3-\alpha}\} \qquad s \ge 1\end{aligned} \tag{3}$$

prob [zero rainfall]

$$\begin{aligned}&= \exp\left\{-\lambda T - \left[\frac{\lambda\nu}{\phi(\alpha-1)}\left(1 + \phi(\kappa+\phi) - \tfrac{1}{4}\phi(\kappa+\phi)\right.\right.\right. \\ &\quad \left.\left.\cdot(\kappa+4\phi) + \frac{\phi(\kappa+\phi)(4\kappa^2+27\kappa\phi+72\phi^2)}{72}\right)\right] \\ &\quad + \frac{\lambda\nu}{(\alpha-1)(\kappa+\phi)}\left(1-\kappa-\phi+\tfrac{3}{2}\kappa\phi+\phi^2+\frac{\kappa^2}{2}\right) \\ &\quad + \frac{\lambda\nu}{(\alpha-1)(\kappa+\phi)}\left(\frac{\nu}{\nu+(\kappa+\phi)T}\right)^{\alpha-1} \\ &\quad \left.\cdot\frac{\kappa}{\phi}\left(1-\kappa-\phi+\tfrac{3}{2}\kappa\phi+\phi^2+\frac{\kappa^2}{2}\right)\right\}\end{aligned} \tag{4}$$

where

$$k_1 = \left(2\lambda\mu_C E^2[x] + \frac{\lambda\mu_C\kappa\phi E^2[x]}{\phi^2-1}\right)\left(\frac{\nu^\alpha}{\alpha-1}\right)$$

$$k_2 = \frac{\lambda\mu_C\kappa E^2[x]}{\phi^2-1}\,\frac{\nu^\alpha}{\alpha-1}$$

The expected duration of a storm is estimated by

$$\begin{aligned}\mu_T \approx \left(\phi\frac{\nu}{\alpha-1}\right)^{-1}\left(1 + \phi(\kappa+\phi) - \frac{\phi(\kappa+\phi)(\kappa+4\phi)}{4}\right. \\ \left. + \frac{\phi(\kappa+\phi)(4\kappa^2+27\kappa\phi+72\phi^2)}{72}\right)\end{aligned} \tag{5}$$

3. Description of Data

For this study, two sets of data are analyzed: one from the Arno basin in central Italy and the other from Paducah, Kentucky. Previously, this model was used to analyze rainfall statistics for the Arno basin in an aggregation mode [*Islam et al.*, 1990]. The emphasis on this paper is in quantifying the ability of the model to disaggregate rainfall statistics. Thus analysis of rainfall data from the Arno basin would allow us to compare the adequacy of the model in aggregation and disaggregation modes.

In a recent study, *Hawk and Eagleson* [1992] analyzed over 70 rain gages from the United States to estimate parameters for the MBRPM and found that only the estimated parameters for Paducah were inconsistent with surrounding station values. They attributed this inconsistency to erroneous data entry for Paducah. Here, we plan to explore this inconsistency with the Paducah data further. We speculate that apparent inconsistency in the *Hawk and Eagleson* [1992] analysis for Paducah might have resulted from the nonunique nature of nonlinear parameter estimation procedure. In fact, we show in the following sections that our procedure satisfactorily reproduces historical statistics for Paducah in both aggregation and disaggregation modes and provides consistent parameter values.

The rain gage from central Italy is located in Borgo, in the Arno River basin. The Arno basin is composed of a cluster of sediment-filled lakes that are connected by steep gorges. Approximately 7000 km^2 of the basin is bounded by the Apennine Mountain range. Generally, elevation increases eastward away from the Mediterranean coast. The complex climatological phenomena observed in the Mediterranean basin are significantly affected by the intensity and distribution of the large-scale pressure systems over the area. These systems interact with the local topography and are strongly influenced by the supply of latent energy from the sea. An important aspect of Mediterranean meteorology is the difference in air and sea temperatures. Mediterranean water is warmer than Atlantic water throughout the year. The presence of island barriers in the Mediterranean serves as a precondition for strong cyclogenesis causing most rainfall over the Arno basin between late fall and early spring. November being the wettest month. The summer months, especially July, are the driest owing to the dominance of the Azores high-pressure cell. Twenty-four years (January 1, 1962, to December 31, 1985) of hourly precipitation data are used from this rain gage.

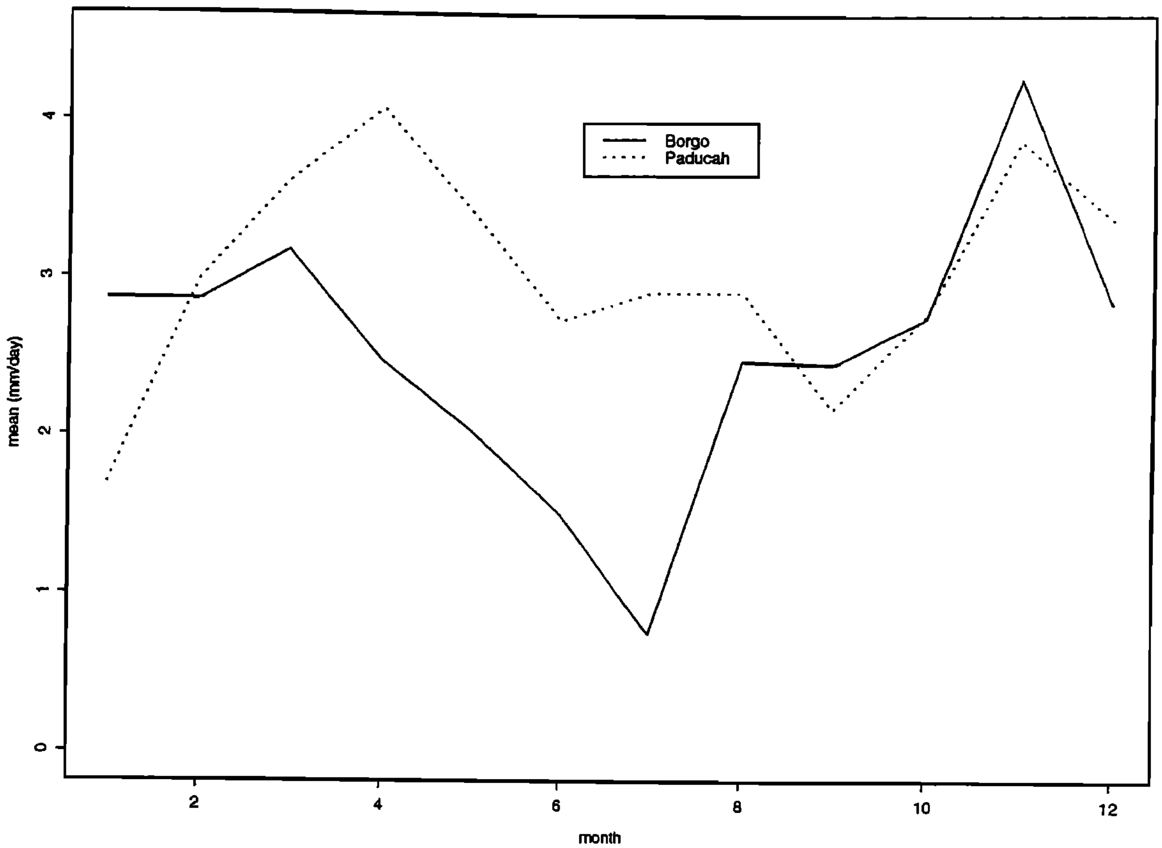

Figure 1a. Seasonal variation of mean for the historical daily rainfall sequence for Borgo and Paducah.

Paducah is a small town in the midwestern United States located near the border of Missouri and Kentucky in a region consisting of mostly rolling plains and grassland. Although Paducah experiences more rainfall than Borgo in spring, the probability of zero rainfall is consistently higher in Paducah for the entire year. January is the driest month in this Midwestern region, while April is the wettest month possibly because of the northern progression of southern storm fronts. Twenty years (January 1, 1970, to December 31, 1989) of hourly precipitation data are used from this station.

Figures 1a, 1b, 2, and 3 show the monthly variation of mean, standard deviation, coefficient of variation, and zero rainfall probability, respectively, for the historical daily rainfall sequence for both Borgo and Paducah. A quick analysis of these figures reveals that although yearly mean rainfall is comparable in these two stations, there are significant differences in other rainfall characteristics. For example, Paducah has consistently higher probability of zero rainfall for the entire year, implying that on average, the storms at Paducah produce more rainfall per storm. Standard deviation of rainfall is more pronounced in Paducah compared with Borgo across all seasons. However, the coefficient of variation is comparable for both stations.

4. Parameter Estimation and Measures of Goodness of Fit

The modified Bartlett-Lewis rectangular pulses rainfall model described in section 2 has six parameters: λ, $E[x]$, α, ν, κ, and ϕ. The parameters are estimated using the method of moments [*Valdez et al.*, 1985; *Islam et al.*, 1990]. We have four theoretical expressions (mean, variance, autocorrelation, and probability of zero rainfall); however, (1) through (4) involve a combination of six model parameters, implying that there are six unknowns and four equations. Except for the mean, these moments are nonlinear functions of accumulation level T. Therefore a combination of these four historical statistics at six accumulation levels would produce 24 equations. As was mentioned before, mean rainfall is a linear function of the accumulation level T, and therefore it cannot be used more than once. Thus we are left with 19 equations and six unknowns. However, we use six equations to estimate six model parameters using method of moments. Estimates of various combinations of first- and second-order statistics from historical precipitation time series may be equated to their theoretical expressions. This results in a set of six highly nonlinear equations with six unknowns. A minimum least squares technique is employed to obtain an estimate of the six model parameters. Let $F(\mathbf{X})$ be the set of nonlinear equations in parameter vector $\mathbf{X}$. In theory, that vector should be equal to the moments of the observation vector Θ. Then,

$$F(\mathbf{X}) - \Theta = 0 \tag{6}$$

where $F(\mathbf{X})$ is the best estimate of Θ. To eliminate the bias due to orders of magnitude difference in the value of Θ, it is convenient to normalize every $F(\mathbf{X})$ by the corresponding Θ value. Now the solution of (6) may be obtained through a simple unconstrained nonlinear minimization:

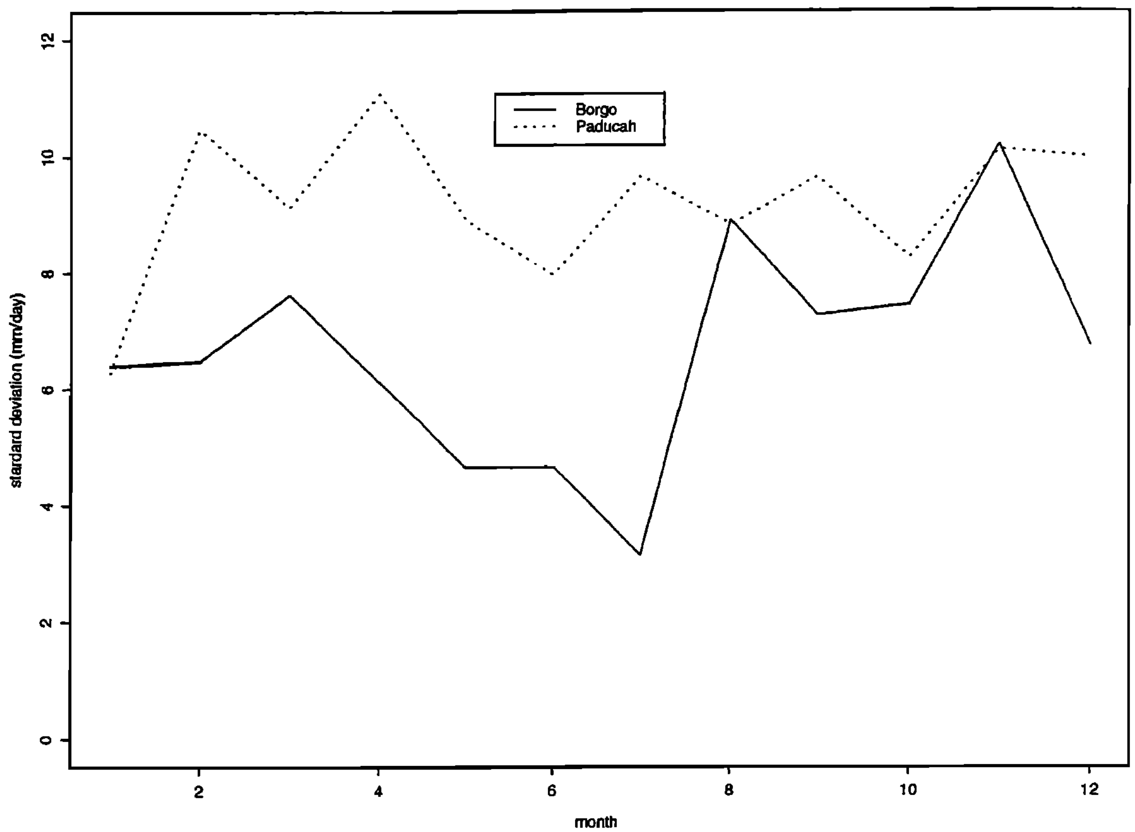

Figure 1b. Seasonal variation of standard deviation for the historical daily rainfall sequence for Borgo and Paducah.

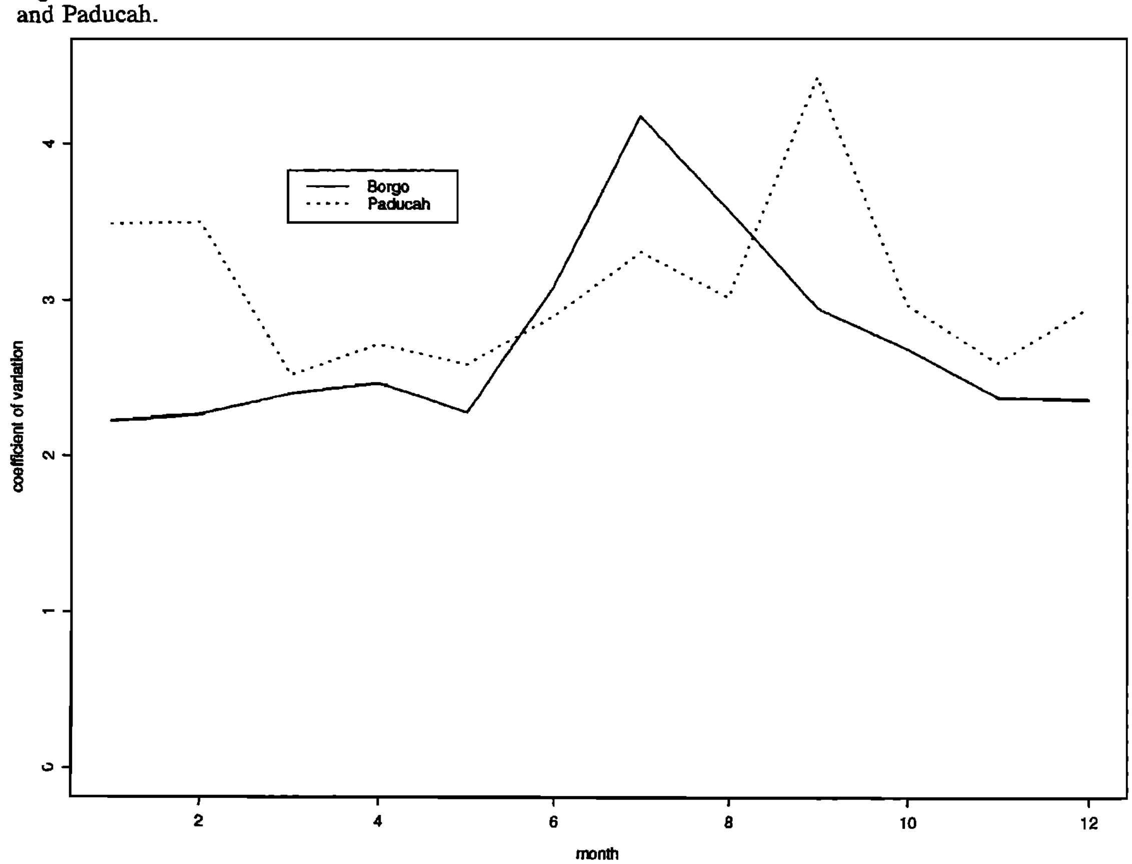

Figure 2. Seasonal variation of coefficient of deviation for the historical daily rainfall sequence for Borgo and Paducah.

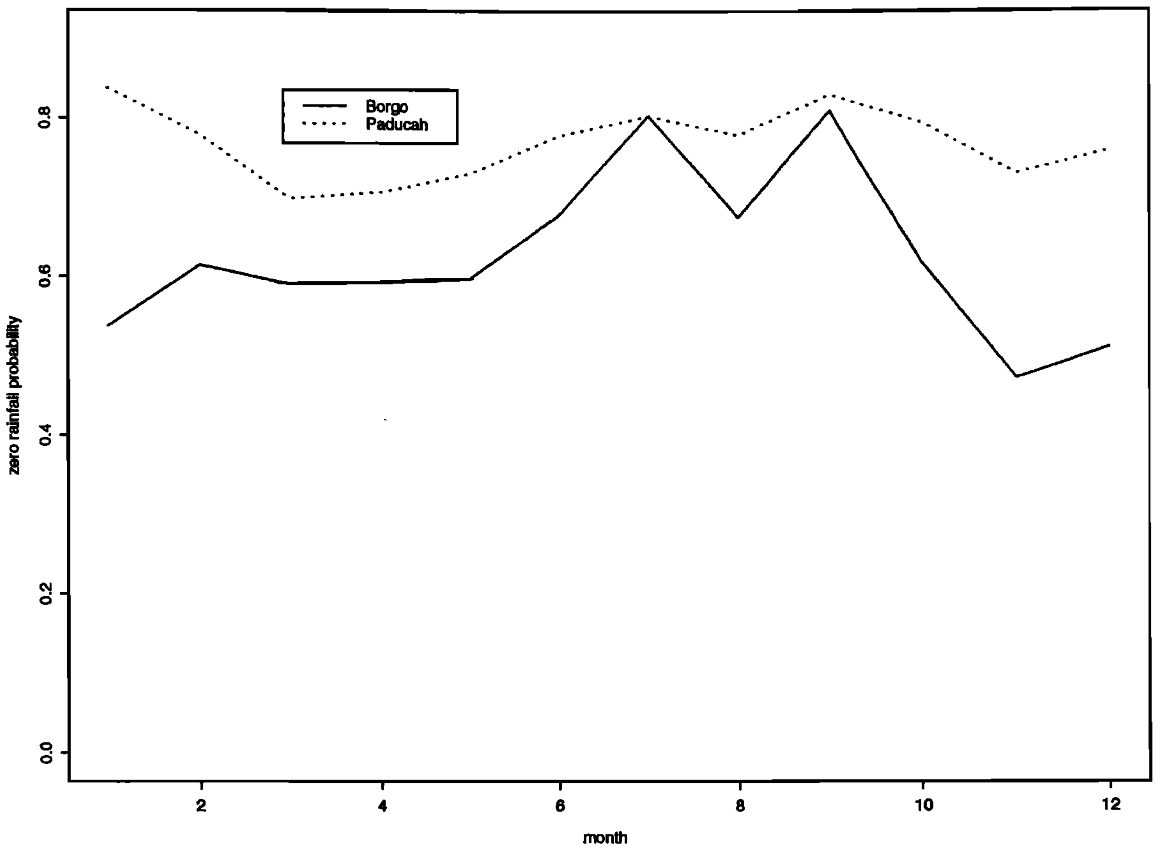

Figure 3. Seasonal variation of zero rainfall probability for the historical daily rainfall sequence for Borgo and Paducah.

$$\min_{X}\left[\left(\frac{F_1(\mathbf{X})}{\Theta_1}-1\right)^2+\left(\frac{F_2(\mathbf{X})}{\Theta_2}-1\right)^2+\cdots+\left(\frac{F_i(\mathbf{X})}{\Theta_i}-1\right)^2+\cdots\right] \quad (7)$$

Parameters are estimated for each month, assuming local stationarity within the month. Tables 1, 2, and 3 show historical and model-reproduced statistics for the cumulative precipitation in Borgo, Italy, at different levels of accumulation for both the aggregation (Table 2) and disaggregation (Table 3) procedure. Clearly, in the aggregation process, finer-scale (1 and 2 hours) statistics are reproduced better than coarser-scale (24 and 48 hours) statistics primarily because we have used some of the finer-scale moments to estimate the parameters. Conversely, in the disaggregation process, coarser-scale statistics are reproduced better. It is worthwhile to note here that only six moments are used in the estimation of parameters. Nevertheless, it appears that at least qualitatively observed statistics at all levels of accumulation are preserved in a satisfactory manner for aggregation as well as disaggregation. Tables 4–6 show similar comparisons for Paducah.

To quantify the adequacy of disaggregation and aggregation procedure in reproducing the historical statistics, we introduce a set of new measures of goodness of fit that accounts for the error between observed and model-fitted quantities for different levels of accumulation. Since mean is a linear function of parameters at all accumulation levels, it is usually reproduced fairly well by the model under almost all circumstances. Therefore we use the other three moments to measure the goodness of fit. These measures are defined as

$$F_1=\frac{1}{N_L}\sum_{i=1}^{N_L}[\ln(\mathrm{Var})_{\mathrm{fitted}}-\ln(\mathrm{Var})_{\mathrm{observed}}]^2 \quad (8)$$

where i = 1-, 2-, 6-, 12-, 24-, and 48-hour accumulation level. N_L refers to the total number of accumulation levels, which is six in our case:

$$F_2=\frac{1}{3N_L}\sum_{i=1}^{N_L}\sum_{j=1}^{3}[(\mathrm{corr}_{ij})_{\mathrm{fitted}}-(\mathrm{corr}_{ij})_{\mathrm{observed}}]^2 \quad (9)$$

where j = 1, 2, and 3 refer to autocorrelation at corresponding lags.

$$F_3=\frac{1}{N_L}\sum_{i=1}^{N_L}[(\%\mathrm{dry})_{\mathrm{fitted}}-(\%\mathrm{dry})_{\mathrm{observed}}]^2 \quad (10)$$

where F_1, F_2, and F_3 refer to measures of goodness of fit for variance, autocorrelation, and probability of zero rainfall, respectively. The measures of goodness of fit for variance, autocorrelation, and probability of zero rainfall are presented in Figures 4, 5, and 6, respectively, for Borgo and Paducah. Figure 4 shows that the historical variance is reproduced fairly well in both locations. It is interesting to note that the fit is extremely good (less than 5% error) for the months of highest variability (November in Borgo and April

Table 1. Historical Statistics of Cumulative Precipitation at Various Levels: Borgo, Italy, November 1–30, 1962–1985

Hour	Mean, mm	Variance, mm^2	Corr(1)	Corr(2)	Corr(3)	Prob(0)
1	0.18	0.81	0.58	0.41	0.33	0.83
2	0.36	2.57	0.53	0.34	0.24	0.80
6	1.07	14.81	0.44	0.18	0.06	0.70
12	2.14	43.59	0.28	0.04	0.03	0.62
24	4.28	108.36	0.18	−0.01	0.01	0.47
48	8.58	242.48	0.12	0.05	0.00	0.30

Table 2. Model-Estimated Statistics of Cumulative Precipitation by Aggregation: Borgo, Italy, November 1–30, 1962–1985

Hour	Mean, mm	Variance, mm^2	Corr(1)	Corr(2)	Corr(3)	Prob(0)
1	0.18*	0.80*	0.62*	0.33	0.21	0.83*
2	0.36	2.58*	0.46	0.19	0.11	0.81*
6	1.07	13.41	0.27	0.09	0.05	0.73
12	2.14	33.96	0.20	0.05	0.02	0.63
24	4.27	81.14	0.13	0.02	0.01	0.47
48	8.54	183.64	0.08	0.00	0.00	0.27

Estimated parameters are $\lambda = 0.0235$ hours, $\nu = 5.3957$, $\alpha = 6.4481$, $E[x] = 2.6615$ mm h^{-1}, $\phi = 0.1186$, and $\kappa = 0.2221$.
*Moments used for parameter estimation.

Table 3. Model-Estimated Statistics of Cumulative Precipitation by Disaggregation: Borgo, Italy, November 1–30, 1962–1985

Hour	Mean, mm	Variance, mm^2	Corr(1)	Corr(2)	Corr(3)	Prob(0)
1	0.18	1.06	0.64	0.35	0.21	0.74
2	0.36	3.47	0.47	0.18	0.09	0.73
6	1.07	18.18	0.25	0.07	0.04	0.67
12	2.13	45.25	0.17	0.05	0.03	0.60
24	4.27*	106.13*	0.13*	0.03	0.02	0.47*
48	8.54	240.28*	0.10	0.01	0.00	0.30*

Estimated parameters are $\lambda = 0.0188$ hours, $\nu = 5.8191$, $\alpha = 5.8878$, $E[x] = 3.6440$ mm h^{-1}, $\phi = 0.0708$, and $\kappa = 0.0836$.
*Moments used for parameter estimation.

Table 4. Historical Statistics of Cumulative Precipitation at Various Levels: Paducah, Kentucky, November 1–30, 1970–1989

Hour	Mean, mm	Variance, mm^2	Corr(1)	Corr(2)	Corr(3)	Prob(0)
1	0.16	0.97	0.46	0.29	0.21	0.96
2	0.32	2.80	0.43	0.22	0.17	0.94
6	0.97	14.19	0.34	0.11	0.10	0.88
12	1.94	38.90	0.25	0.07	0.01	0.82
24	3.88	101.93	0.15	−0.05	0.05	0.73
48	7.76	229.11	0.03	0.07	−0.01	0.55

Table 5. Model-Estimated Statistics of Cumulative Precipitation by Aggregation: Paducah, Kentucky, November 1–30, 1970–1989

Hour	Mean, mm	Variance, mm^2	Corr(1)	Corr(2)	Corr(3)	Prob(0)
1	0.16*	0.96*	0.46*	0.35	0.27	0.95*
2	0.32	2.81*	0.49	0.30	0.18	0.94*
6	0.96	15.56	0.35	0.09	0.03	0.89
12	1.92	42.00	0.21	0.02	0.00	0.83
24	3.84	101.38	0.10	0.00	0.00	0.73
48	7.68	223.35	0.05	0.00	0.00	0.55

Estimated parameters are $\lambda = 0.0100$ hours, $\nu = 4.7806$, $\alpha = 30.5887$, $E[x] = 5.9149$ mm hour, $\phi = 0.0307$, and $\kappa = 0.4905$.
*Moments used for parameter estimation.

Table 6. Model-Estimated Statistics of Cumulative Precipitation by Disaggregation: Paducah, Kentucky, November 1-30, 1970-1989

Hour	Mean, mm	Variance, mm^2	Corr(1)	Corr(2)	Corr(3)	Prob(0)
1	0.16	1.07	0.36	0.23	0.20	0.93
2	0.32	2.91	0.38	0.26	0.19	0.92
6	0.97	14.59	0.37	0.15	0.06	0.87
12	1.94	39.87	0.27	0.05	0.01	0.82
24	3.87*	101.05*	0.15*	0.01	0.00	0.72*
48	7.74	231.87*	0.07	0.00	0.00	0.56*

Estimated parameters are $\lambda = 0.0107$ hour, $\nu = 5.7004$, $\alpha = 30.4587$, $E[x] = 7.3266$ mm h^{-1}, $\phi = 0.0295$, and $\kappa = 0.2840$.
*Moments used for parameter estimation.

in Paducah). Figure 5 shows satisfactory goodness of fit for correlation. The difference between the observed and fitted correlation range between 0.04 and 0.14 with no seasonal or geographical structure. On average, the model captures the probability of zero rainfall for Paducah with less than 3% error (Figure 6) for the entire year. For Borgo, errors are less than 5% except for the month of March, when error goes as high as 10%.

Thus it is clear that in all three goodness of fit measures, the disaggregation procedures perform very well. Now let us compare the performance measures for aggregation and disaggregation. Figures 7 and 8 compare the goodness of fit for variance for Paducah and Borgo, respectively. From Figure 7 it would appear that the disaggregation procedure consistently outperforms the aggregation model in reproducing historical variance; however, Figure 8 shows a rather mixed performance. Similar results (not shown here for brevity) are also obtained for other two goodness of fit measures. Therefore one may reasonably claim that using readily available daily rainfall data, the model can reproduce finer-scale statistics.

Now, to illustrate the physical realism of the estimated parameters in the disaggregation model, let us briefly examine two important model characteristics: arrival time between storms and storm duration. Figure 9 shows the seasonal variability in the arrival time between storms (λ^{-1}); it is evident that storms are more frequent in Borgo than in Paducah throughout the year. Figure 10 shows the average length of storm duration as a function of time of year. On an average, Borgo experiences storms with longer duration. These two model characteristics are consistent with the observed probability of zero rainfall as shown in Figure 3.

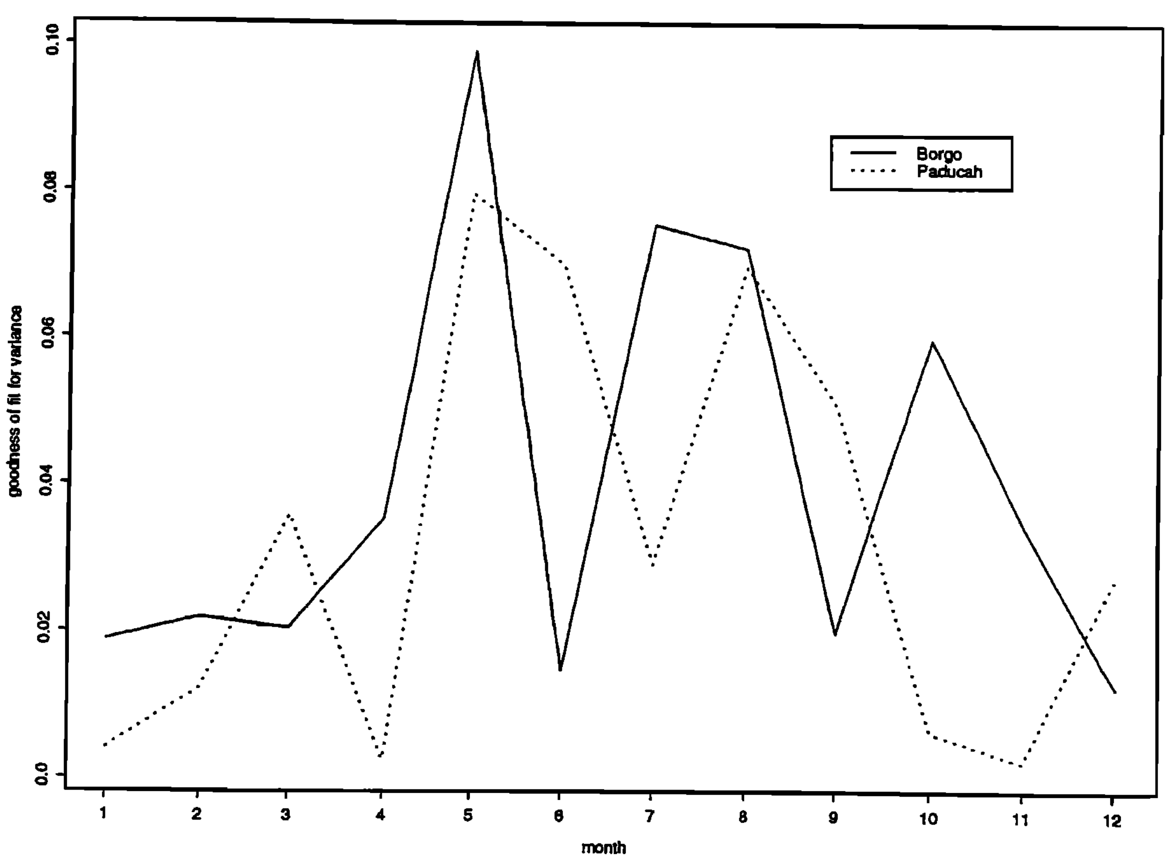

Figure 4. Goodness of fit for variance through the year.

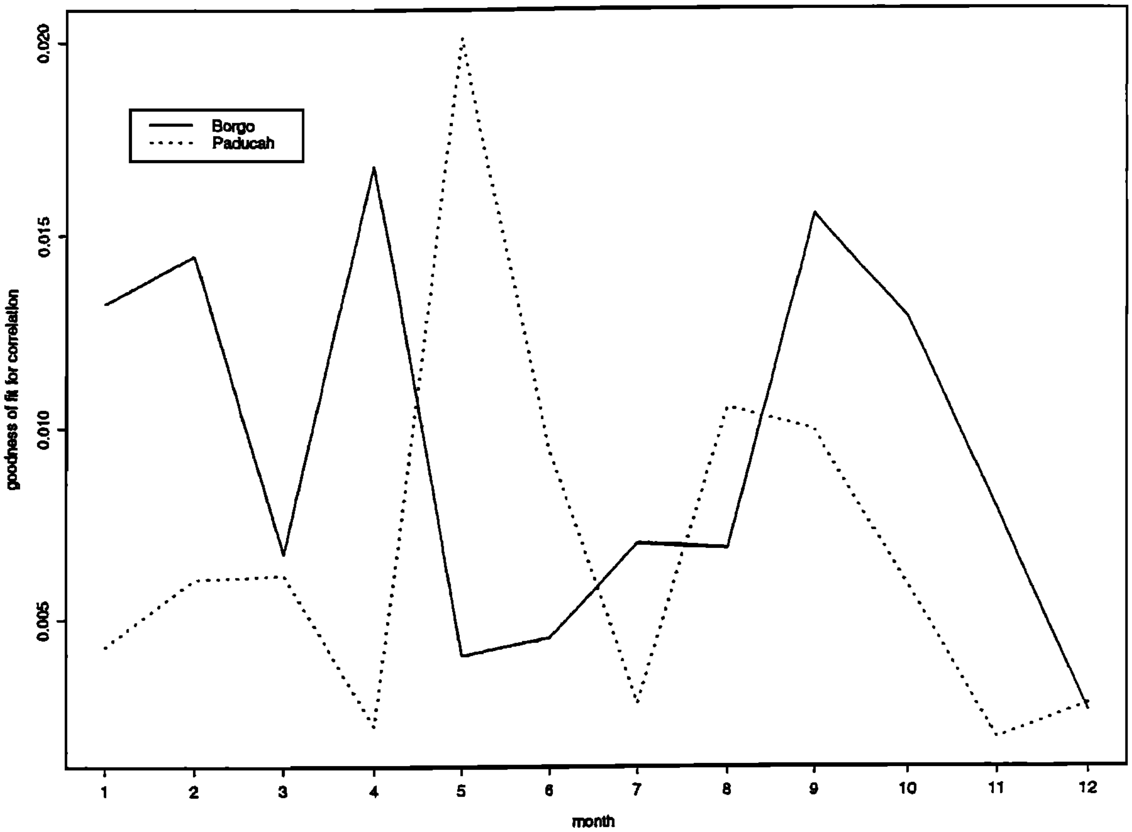

Figure 5. Goodness of fit for autocorrelation through the year.

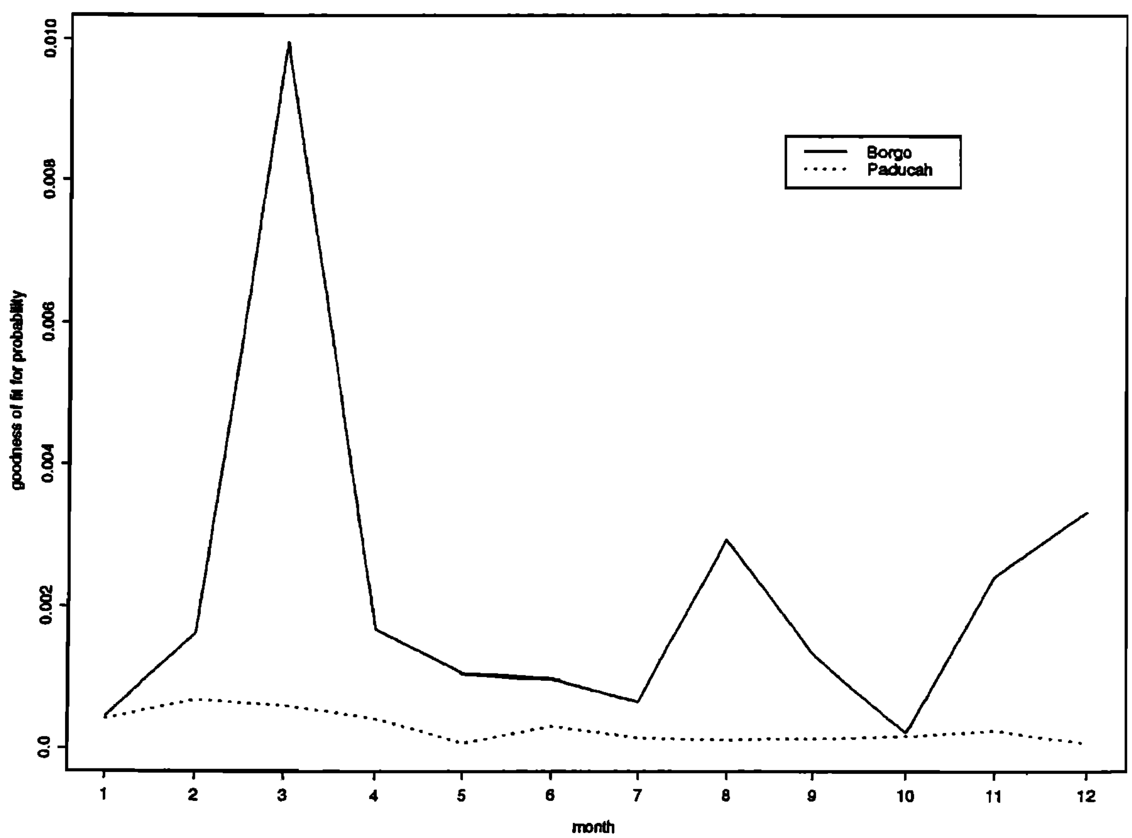

Figure 6. Goodness of fit for probability of zero rainfall through the year.

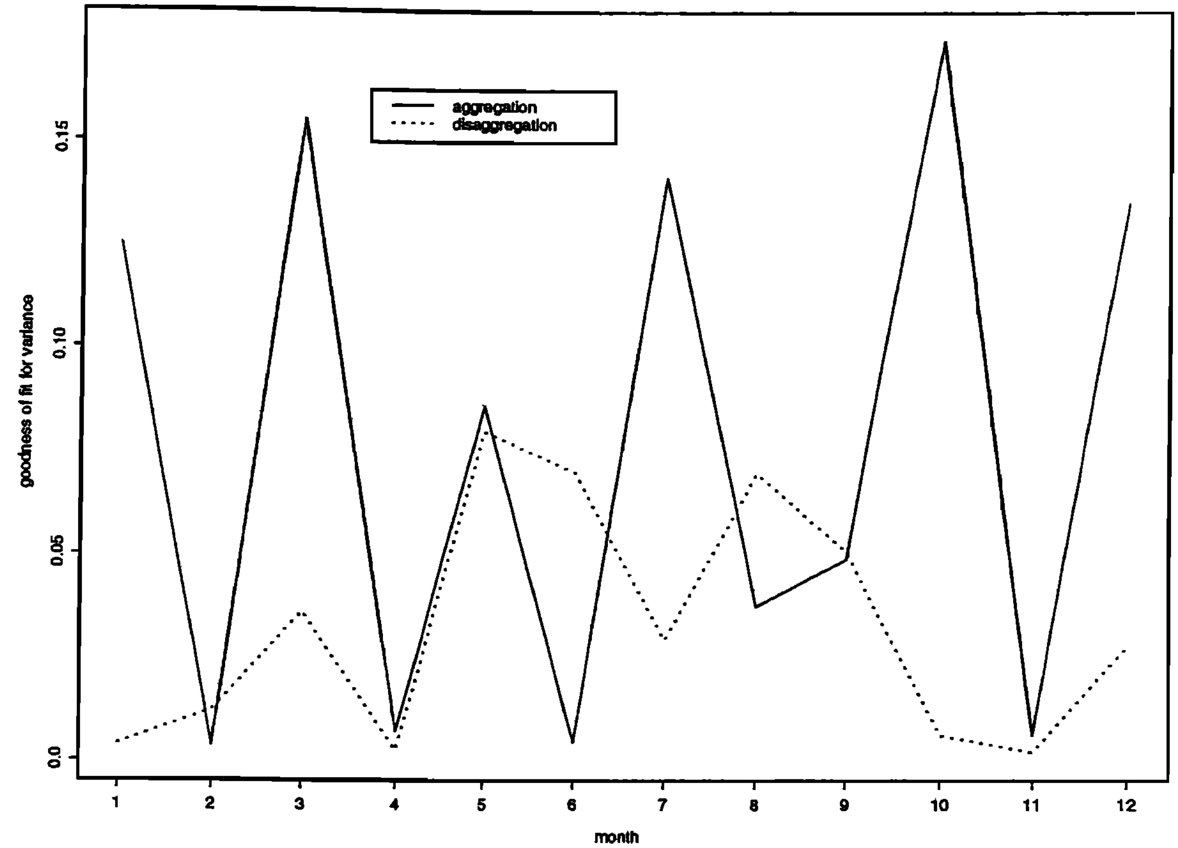

Figure 7. Comparison of goodness of fit for variance between aggregation and disaggregation in Paducah.

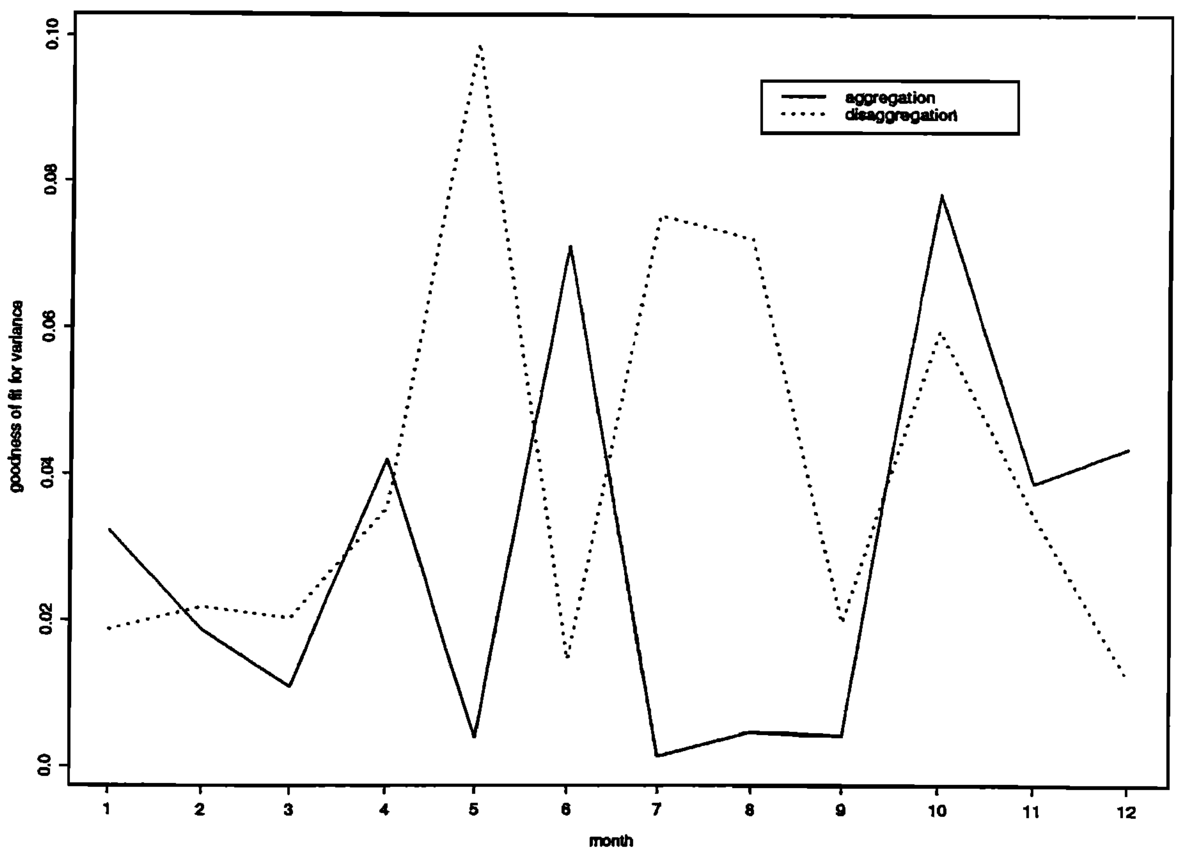

Figure 8. Comparison of goodness of fit for variance between aggregation and disaggregation in Borgo.

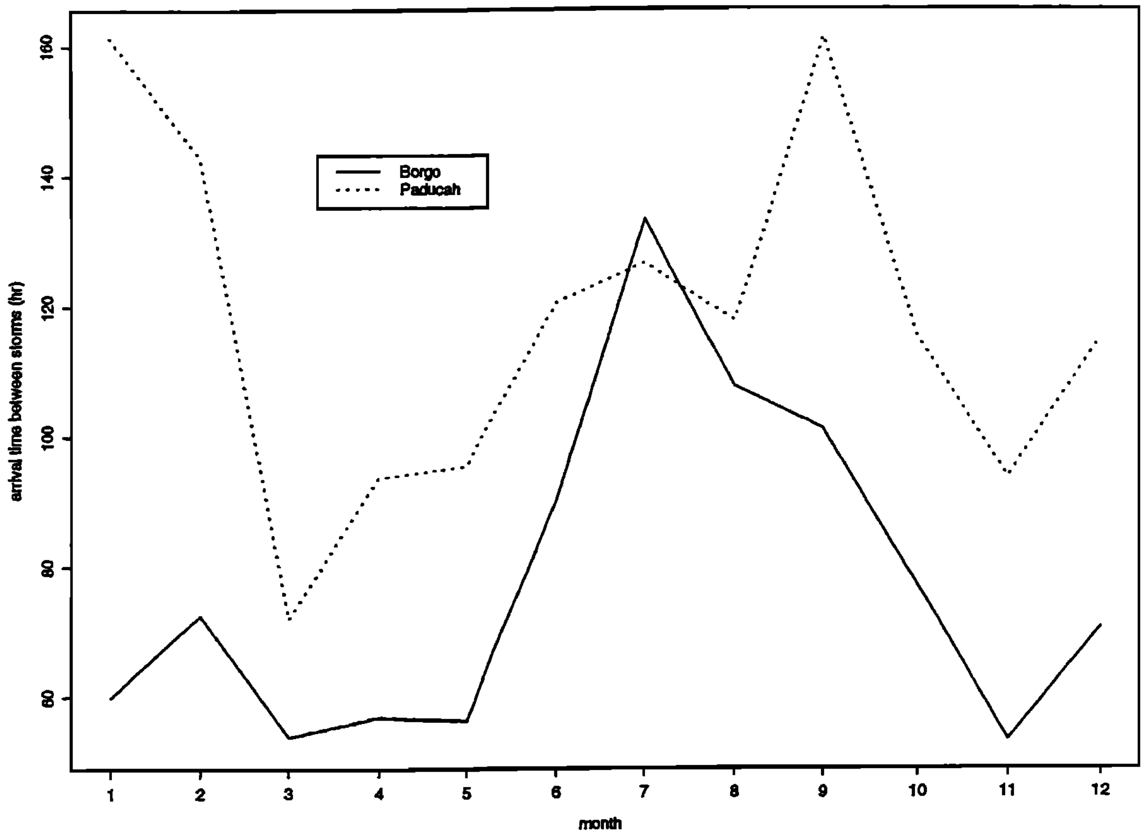

Figure 9. Arrival time between storms (λ^{-1}) through the year for Borgo and Paducah.

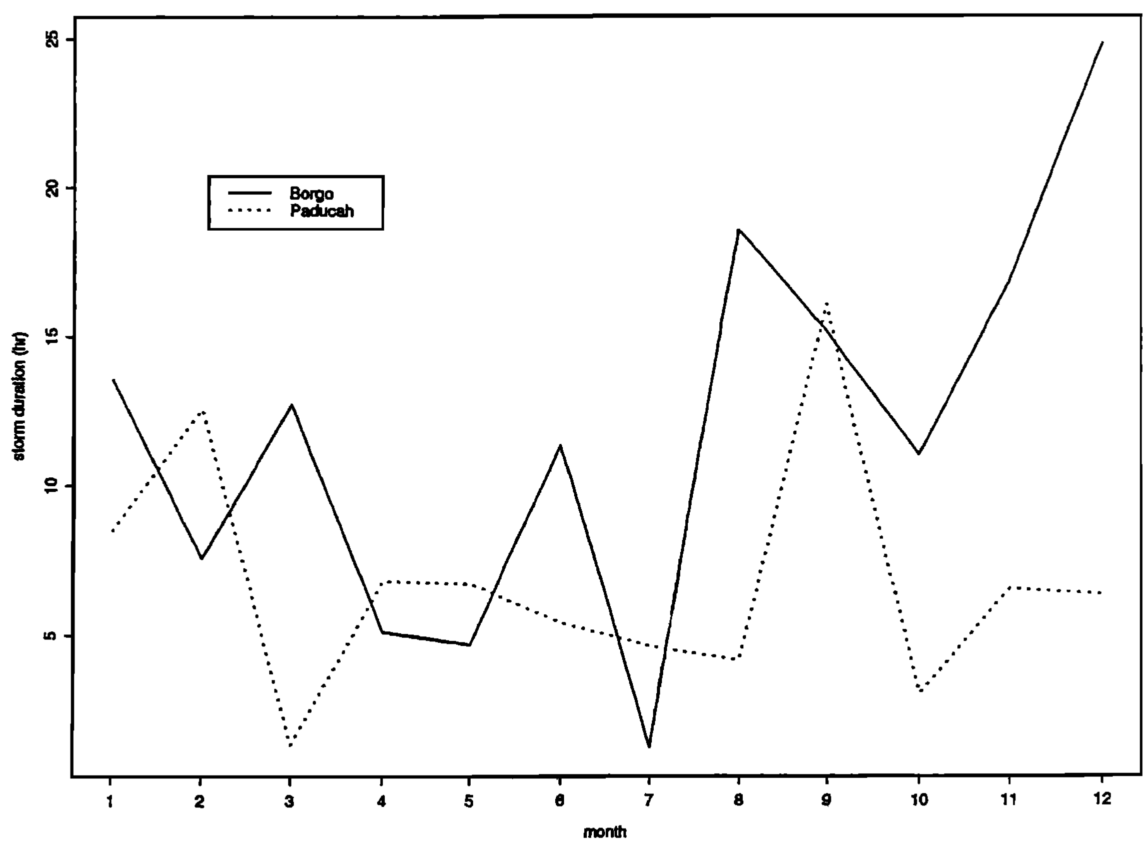

Figure 10. Storm duration through the year for Borgo and Paducah.

Furthermore, for all seasons the model-produced storm durations are within a physically reasonable range. For brevity of presentation we have shown results from two rain gages only; however, analyses of rainfall data from several other stations from Denver, Colorado, and the Arno basin in Italy (E. Elathir, unpublished manuscript, 1994) provide evidence of applicability of this model, in aggregation as well as disaggregation modes, to a variety of rainfall regimes.

5. The Power Spectrum for the Bartlett-Lewis Model

From the analysis of section 4, we find that the modified Bartlett-Lewis model is capable of disaggregating rainfall statistics from a daily timescale to hourly scale. To better understand this self-consistent aspect of the model, we will now look at the power spectral density function of the model. Here, self-consistency implies that even after using information from a coarser timescale, the model is able to reproduce historical statistics at finer timescales. The term power spectral density function is often shortened to power spectrum. The power spectrum is a frequency decomposition of the variance of a process. It reflects the contribution of each frequency to the overall variability of the process. If the power spectrum does not show any preferential frequencies but depicts a tendency to link a wide range of frequencies, one might argue that the model variability is arising from a wide range of frequencies and the model should be able to aggregate and disaggregate rainfall at various temporal scales. Mathematically, the power spectrum is defined as the Fourier transform of the autocovariance function as follows:

$$f^T(\omega) = \frac{1}{\Pi} \int_{-\infty}^{\infty} \gamma_Y^T(s) e^{-i\omega s} \, ds \qquad 0 < \omega < \infty \tag{11}$$

where

$$\gamma_Y^T(s) = \text{Cov}\,[Y^T(t),\ Y^T(t+s)]$$

where $\gamma_Y^T(s)$ is the autocovariance function for an aggregation level T as defined in (3). Realizing that $\gamma_Y^T(s)$ is an even function, we may write (11) as follows:

$$f^T(\omega) = \frac{2}{\Pi} \int_0^{\infty} \gamma_Y^T(s) \cos \omega s \, ds \tag{12}$$

Equation (12) defines the power spectrum for all positive ω for a given level of accumulation T. Equation (12) implies an integration over an infinite time lag for the autocovariance function. For all practical purposes, the autocovariance function goes to zero within a finite number of lags. Thus we approximate (12) by summing it over 50 lags. In all cases we tested, the autocovariance function becomes insignificant within 20 lags. A numerical approximation of (12) may be written as

$$f^T(\omega) = \frac{2}{\Pi} \sum_{k=0}^{50} \gamma_Y^T(k) \cos \omega k \tag{13}$$

It is sometimes useful to use a normalized form of the power spectrum, i.e., (13) is normalized by the variance of the process. Therefore cumulative area under the normalized power spectrum will be unity. An important parameter for the power spectrum function for the modified Bartlett-Lewis model is the shape parameter α. When α approaches to 2, the power spectrum resembles the fractional Gaussian noise process of *Mandelbrot and Wallis* [1968]. On the other hand, for $\alpha > 3$, correlation decays very rapidly.

Figure 11 shows the normalized power spectrum in aggregation and disaggregation modes for hourly accumulation scale. For the power spectrum, parameters used are from Borgo for November (Tables 2 and 3). Clearly, there are no preferential frequencies where the power is concentrated. Instead, the power is distributed across a range of frequencies. However, it is interesting to note that the power spectrum is essentially flat (i.e., the underlying process is a random noise with no temporal structure) for frequencies less than the storm arrival rate. This implies that variability at timescales greater than the storm arrival rate is due to the independent arrival of storm centers. For the analyzed stations, 2 days ($\lambda = 0.0208$ hours^{-1}) is near the temporal scales that separate within-storm structured variability and between-storm independent variability. Figure 12 shows the dependence of power spectrum on the storm arrival rate. For timescales greater than 2 days there is no derivable information because the spectrum essentially resembles a white noise spectrum.

As we can see from Figure 11, the power spectrum for hourly accumulation level in both aggregation and disaggregation models has very similar structures. Thus although aggregation and disaggregation procedures provide different set of estimated parameters, the underlying variability structure is well preserved through both the models. This leads us to speculate that the structure of the model is self-consistent for upscaling (aggregation) or downscaling (disaggregation) of rainfall statistics for the temporal scales between hours and 2 days. Although we have not shown it from the data, given the structure of the power spectrum, it is quite possible that this model will be able to disaggregate rainfall statistics to minute level from daily level. We are currently exploring this possibility using fine-scale rainfall data.

Although the analytical form of the power spectrum is quite complex, on the basis of the power spectrum structure it can be approximated as follows:

$$\begin{aligned} f(\omega) &= f_0 & \omega &< \omega_1 \\ f(\omega) &= f_0(\omega/\omega_1)^{-k} & \omega &\geq \omega_1 \end{aligned} \tag{14}$$

In words, the power spectrum is approximated to have a horizontal part connected with an inclined line having a constant slope k. The transition between the horizontal part and the inclined one is controlled by the timescale T_1 ($= 1/\omega_1$). Based on this approximation, to capture the whole spectrum, it is sufficient to have information from timescales $T_2 < T_1$, such that the slope k can be estimated. Indeed, as was shown through the goodness of fit tests (section 4), statistics at finer scales can be inferred from daily level statistics. The key feature that makes disaggregation and aggregation possible is the power law dependence of the power spectrum for timescales smaller than T_2. Therefore one can speculate that the model will perform poorly in disaggregating rainfall when calibrated with information at timescales larger than T_1.

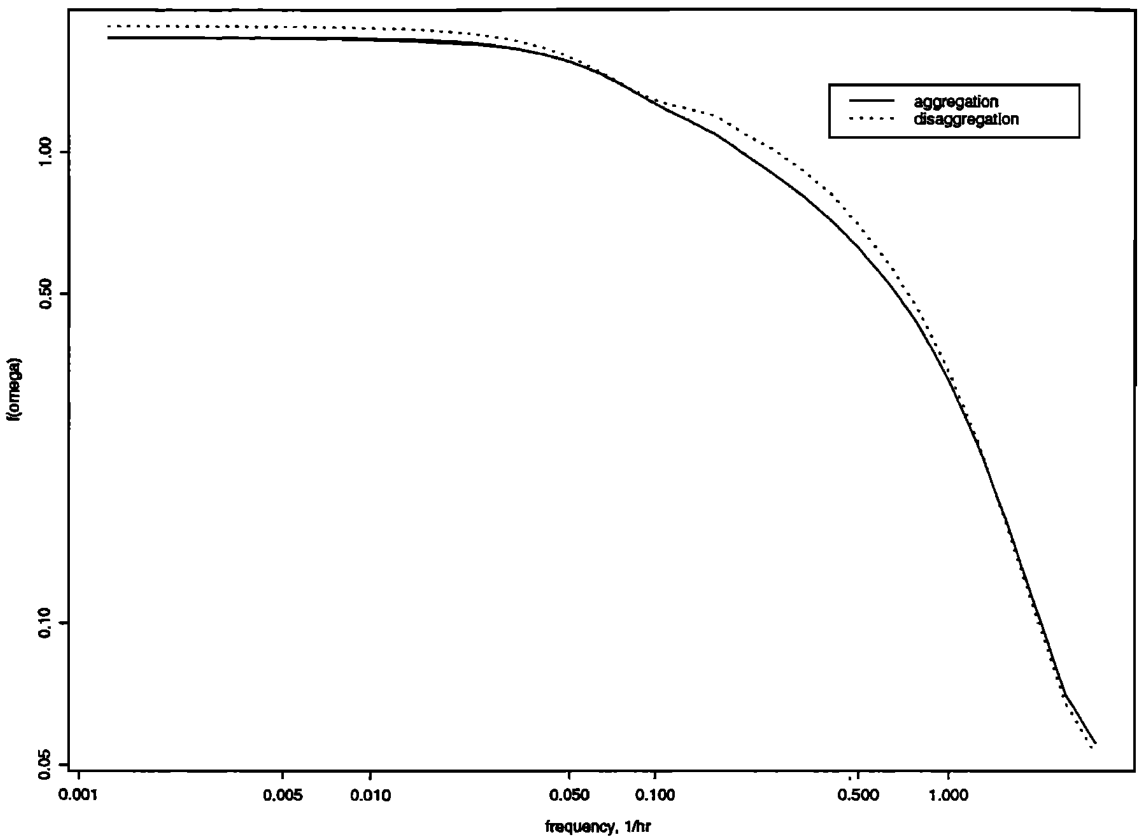

Figure 11. The normalized power spectrum in aggregation and disaggregation modes for hourly level in November for Borgo.

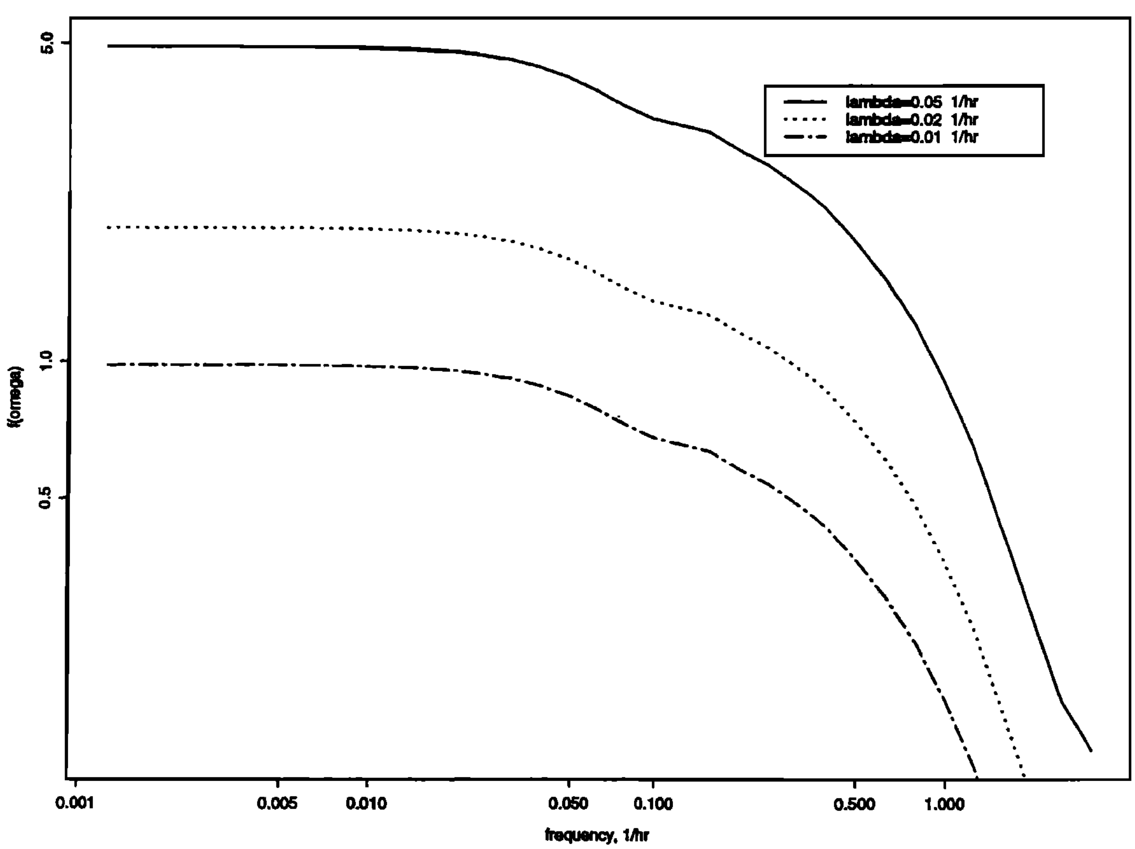

Figure 12. The dependence of power spectrum on the storm arrival rate.

6. Concluding Remarks

Finer-timescale rainfall data is required for a variety of hydrologic applications. However, most of the routinely collected rainfall data are archived on a daily timescale. In this paper, we use a point process rainfall model to aggregate and disaggregate rainfall statistics at different timescales.

To quantify the adequacy of aggregation and disaggregation procedure in reproducing historical statistics, we introduce a set of new measures of goodness of fit for variance, autocorrelation, and probability of zero rainfall at different levels of accumulation. Based on the results of two rainfall stations, we find that using readily available daily rainfall data, the modified Bartlett-Lewis Rectangular Pulses Model can reproduce finer-timescale rainfall statistics.

An explanation based on the structure of the power spectrum is proposed for the adequacy of the aggregation and disaggregation properties of the model. It appears that approximately a 2-day timescale separates the interstorm and intrastorm variabilities. Since the power spectrum is essentially flat for timescales greater than 2 days, finer-scale statistics cannot be obtained from aggregation levels greater than 2 days. However, the model can disaggregate to very fine scale statistics using the information from daily accumulation levels. This characteristic behavior of the model is related to the power law dependence of the power spectrum for timescales smaller than 2 days.

A natural extension of this work will be to investigate the similarities between the observed and model-produced power spectra. If the observed spectrum can be approximated by the proposed three-parameter model, then it can be used to aggregate and disaggregate rainfall for timescales shorter than 2 days.

Acknowledgments. This work was supported in part by a grant from the National Institute for Global Environmental Change.

References

Econopouly, T. W., D. R. Davis, and D. A. Woolhiser, Parameter transferability for a daily rainfall disaggregation model, *J. Hydrol.*, *118*, 209–228, 1990.

Hawk, K., and P. S. Eagleson, Climatology of station storm rainfall in the continental United States: Parameters of the Bartlett Lewis and Poisson rectangular pulses model, *Rep. 336*, Ralph M. Parsons Lab., Mass. Inst. of Technol., Cambridge, 1992.

Hershenhorn, J., and D. A. Woolhiser, Disaggregation of daily rainfall, *J. Hydrol.*, *95*, 299–322, 1987.

Islam, S., D. Entekhabi, R. L. Bras, and I. Rodriguez-Iturbe, Parameter estimation and sensitivity analysis for the modified Bartlett-Lewis rectangular pulse model of rainfall, *J. Geophys. Res.*, *95*(D3), 2093–2100, 1990.

Mandelbrot, B. B., and J. R. Wallis, Noah, Joseph, and operational hydrology, *Water Resour. Res.*, *4*(5), 909–918, 1968.

Onof, C., and H. S. Wheater, Modelling of British rainfall using a random parameter Bartlett-Lewis rectangular pulses model, *J. Hydrol.*, *149*, 67–95, 1993.

Rodriguez-Iturbe, I., D. R. Cox, and V. Isham, Some models for rainfall based on stochastic point processes, *Proc. R. Soc. London 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 A*, *417*, 283–298, 1988.

Srikanthan, R., and T. A. McMahon, Stochastic generation of rainfall and of evaporation data, *Aust. Water Res. Counc. Tech. Pap. 894*, 301 pp., Australian Gov. Publ. Service, Canberra, A.C.T., 1985.

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

Z. Bo and S. Islam, Department of Civil and Environmental Engineering, P.O. Box 210071, University of Cincinnati, Cincinnati, OH 45221-0071.

E. A. B. Eltahir, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

(Received May 16, 1994; revised August 1, 1994;
accepted August 5, 1994.)