WATER RESOURCES RESEARCH, VOL. 22, NO. 6, PAGES 965–978, JUNE 1986 #393

Seasonal and Regional Variability of Parameters for Stochastic Daily Precipitation Models: South Dakota, U. S. A.

D. A. WOOLHISER

U.S. Department of Agriculture, Agricultural Research Service, Tucson, Arizona

JOSÉ ROLDÁN

Catedra de Hidraulica General y Agricola, Escuela Tecnica Superior de Ingenieros, Agronomos, Universidad de Cordoba, Cordoba, Spain

Fourier series are used to describe the seasonal variation of the five parameters for a stochastic model of daily precipitation utilizing a first-order Markov chain for the occurrence process and a mixed exponential distribution for the daily precipitation amounts (MCME model). Spatial variability of the means of each parameter for 16 stations in South Dakota has been illustrated by mapping isopleths. MCME parameters for 4 stations not included in the analysis are more closely described by the arithmetic mean of parameters for the 6 nearest stations than by using parameters for the nearest neighboring station or parameters estimated by spline fitting or linear interpolation. However, MCME parameters estimated by all interpolation methods were significantly different from parameters identified for each of the four stations by maximum likelihood techniques. The principal source of this spatial variability at distances of on the order of 100 km is data inconsistency due to methodological differences affecting small precipitation amounts and apparently related to observation time. Sampling error, possible parameter identifiability problems, and real differences in the precipitation regime on a scale smaller than the station spacing also contribute to the observed variability.

Oh Dakota land,
Sweet Dakota land,
As on thy fiery soil I stand,
I look across the plains,
And wonder why it never rains,
'Til Gabriel blows his trumpet sound,
And says the rain's just gone around ...
(Chorus)

"Dakota Land," early settlers' folk song

INTRODUCTION

Agriculturalists all over the world, and particularly those who live in the extensive semiarid regions, have long recognized the importance of seasonal and annual variability of precipitation to their livelihood. Information on rainfall probabilities is vital for the design of water supply and supplemental irrigation schemes and the evaluation of alternative cropping and of soil and water management plans. Such information can also be beneficial in determining the best adapted plant species and the optimum time of seeding to reestablish vegetation on deteriorated rangelands. Although rather long precipitation records are frequently available in many countries, little use is made of this information because of the unwieldy nature of the records. Many efforts have been made to summarize precipitation records into more useful forms. For example, in the United States, regional research groups have published extensive tables of Markov chain parameters and tabulated values of distribution functions for daily rainfall [cf. *Feyerherm et al.*, 1965; *Barger et al.*, 1959].

Although these publications are more useful than the raw data, in most circumstances they are not well adapted to provide input for simulation models, nor do they provide a sufficiently concise description of the precipitation climatology of a region to improve our ability to differentiate between different rainfall regimes or to group similar regimes.

In an effort to provide more concise models of daily precipitation, several investigators have proposed stochastic models describing both precipitation occurrence and the distribution of precipitation amounts at a point in space [*Jones et al.*, 1972; *Smith and Schreiber*, 1973, 1974; *Todorovic and Woolhiser*, 1975; *Haan et al.*, 1976; *Buishand*, 1977; *Katz*, 1977; *Woolhiser and Pegram*, 1979; *Stern*, 1980*a*, *b*].

The seasonal variation in stochastic model parameters has long been recognized and is usually accounted for by estimating the parameters for time periods of a few days to a few months and assuming that the parameters are constant within the period but take discrete jumps at the end of the period [cf. *Richardson and Wright*, 1984]. Several investigators have fitted Fourier series to model parameters [*Feyerham and Bark*, 1965; *Buishand*, 1977; *Coe and Stern*, 1982; *Woolhiser et al.*, 1973]. *Jones et al.* [1972] and *Stern* [1980*a*] fit polynomial curves to first-order Markov chain parameters.

Woolhiser and Pegram [1979] used direct numerical maximum likelihood estimates of Fourier coefficients to describe the seasonal variations of parameters in a stochastic model of daily precipitation. They demonstrated the technique using a first-order Markov chain as the occurrence process and a mixed exponential distribution for the daily precipitation, and they suggested that the means, amplitudes, and phase angles of the Fourier series for each parameter could be mapped to provide a parsimonious regionalized model of the point precipitation process.

Roldán and Woolhiser [1982] and *Woolhiser and Roldán* [1982] compared the Markov chain-mixed exponential (MCME) model with several alternatives for five widely-scattered stations in the United States. According to the Akaike information criterion [*Akaike*, 1974], the MCME was superior to the alternatives studied, including chain dependent models, i.e., where the distribution of precipitation on a wet day is dependent on the state of the previous day.

Paper number 5W4266.
0043-1397/86/005W-4266$05.00

The objective of this paper is to examine the spatial characteristics of the Fourier coefficients for each parameter in the MCME model using data for the state of South Dakota, as an example. We also examine alternative interpolation techniques for estimating Fourier coefficients at points between stations.

STOCHASTIC DAILY PRECIPITATION MODEL

The precipitation occurrence process X_t is described by a first-order Markov chain, with two states defined by the transition probabilities

$$p_{ij}(n) = P(X_n = j \mid X_{n-1} = i) \qquad i, j = 0, 1 \qquad n = 1, 2, \cdots 365 \tag{1}$$

where state 0 signifies a dry day and state 1 a wet day, and

$$p_{i1}(n) = 1 - p_{i0}(n) \qquad i = 0, 1 \tag{2}$$

Let Y_t be the amount of precipitation that falls on day t when $X_t = 1$. We assume that Y_t is serially independent and is independent of X_{t-1}. This means that there is dependence on precipitation occurrence from day to day but that the amount of precipitation (given a wet day) is independent of previous occurrences and amounts. The assumption of independence between the amounts of rainfall on successive days leads to significant simplifications in the model structure and has been used by several previous investigators [cf. *Coe and Stern*, 1982; *Richardson and Wright*, 1984; *Stern*, 1980*a*], although small but significant correlations have been found [*Buishand*, 1977].

Let the random variable $U_t = Y_t - T$ be distributed as a mixed exponential with probability density function:

$$f_n(u) = \frac{\alpha(n)}{\beta(n)} \exp\left(-\frac{u}{\beta(n)}\right) + \frac{[1-\alpha(n)]}{\delta(n)} \exp\left(-\frac{u}{\delta(n)}\right) \tag{3}$$

where $0 \le u \le \infty$, T is a threshold (normally 0.25 mm or 0.01 inch), $0 \le \alpha(n) \le 1$, $0 < \beta(n) < \delta(n)$, and $n = 1, 2, \cdots 365$.

The mixed exponential distribution can be interpreted as the result of a random sample from two exponential distributions where the distribution with the smaller mean $\beta(n)$ is sampled with probability $\alpha(n)$ and the distribution with the larger mean $\delta(n)$ is sampled with probability $(1 - \alpha(n))$.

It has also been shown [*Woolhiser and Pegram*, 1979] that the seasonal variations in each of the five parameters can be described by the polar form of a finite Fourier series:

$$\gamma_i(n) = \gamma_{i0} + \sum_{j=1}^{m_i} \left\{ C_{ij} \sin\left(\frac{2\pi nj}{365} + \phi_{ij}\right)\right\} \tag{4}$$

where $i = 1, 2, \cdots 5$, $\gamma_i(n)$ is the value of the ith parameter on day n, $n = 1, 2, \cdots 365$, m_i = maximum number of harmonics for the ith parameter, γ_{i0} = mean of each parameter, C_{ij} = amplitude, and ϕ_{ij} = phase angle.

The expected value function of the total precipitation in k days $S(k)$ can be written in the following form:

$$\begin{aligned} E[S(k)] &= \sum_{i=1}^{k} E[Y_i \quad X_i] = \sum_{i=1}^{k} E[U_i \quad X_i] + T\sum_{i=1}^{k} E[X_i] \\ &= \sum_{i=1}^{k} [P(X_{i-1} = 0)p_{01}(i) + P(X_{i-1} = 1)p_{11}(i)] \\ &\quad \cdot \{\alpha(i)\beta(i) + [1 - a(i)]\delta(i) + T\} \end{aligned} \tag{5}$$

PARAMETER ESTIMATION

Woolhiser and Pegram [1979] recommended direct numerical maximum likelihood techniques to estimate the Fourier coefficients. *Coe and Stern* [1982] have formulated this problem as a generalized linear model which allows a regression-type approach to be used to fit and test alternative models. In this paper we use a procedure similar to that used by *Woolhiser and Pegram* [1979]. Both occurrence and precipitation depth were handled independently, but the optimization process followed was almost the same in both cases. First of all, parameters of both the Markov Chain and the mixed exponential distribution were computed by maximum likelihood methods for 14-day periods. Maximum likelihood estimates of Markov Chain parameters are easily calculated by computing the observed number of transitions $a_{ij}(n)$ from state i (0 or 1) on day $n - 1$ to state j (0 or 1) on day n. Then

$$p_{00}(k) = \frac{a_{00}(k)}{a_{00}(k) + a_{01}(k)} \tag{6}$$

$$p_{10}(k) = \frac{a_{10}(k)}{a_{10}(k) + a_{11}(k)} \tag{7}$$

where $k = 1, 26$ and $a_{ij}(k)$ refers to the number of transitions occurring within period k.

The maximum likelihood estimates of the parameters of the mixed exponential distribution were obtained by maximizing the log likelihood function:

$$\log L_k = \sum_{j=1}^{N(k)} \left\{ \log \left[\frac{\alpha_k}{\beta_k} \exp(-u_{kj}/\beta_k) + \left(\frac{1-\alpha_k}{\delta_k}\right) \exp(-u_{kj}/\delta_k)\right]\right\} \tag{8}$$

where α_k, β_k, and δ_k are the parameter values for the kth period, $k = 1, 2, \cdots 26$, u_{kj} = the amount of precipitation minus the threshold for the jth wet day in period k, and $N(k)$ is the number of wet days in period k.

The Fourier coefficients for each of the five parameters were first estimated using least squares methods. These coefficients were used as starting values for maximum likelihood estimation of coefficients using a multivariate, unconstrained optimization technique from the IMS (International Mathematical and Statistical) Library, called ZXMIN, based on a paper by *Fletcher* [1972].

The coefficients γ_{i0}, C_{ij}, ϕ_{ij}, $i = 1, 5$; $j = 1, m$ for the MCME were estimated by maximizing the following expressions [*Woolhiser and Pegram*, 1979]:

Dry-dry transitions

$$\log L_d = \sum_{n=1}^{365} \{a_{00}(n) \log p_{00}(n) + a_{01}(n) \log [1 - p_{00}(n)]\} \tag{9}$$

Wet-dry transitions

$$\log L_w = \sum_{n=1}^{365} \{a_{10}(n) \log p_{10}(n) + a_{11}(n) \log [1 - p_{10}(n)]\} \tag{10}$$

where $a_{ij}(n)$ is the observed number of transitions from state i(0 or 1) on day $n - 1$ to state j(0 or 1) on day n. Mixed exponential distribution:

$$\log L_{ME} = \sum_{n=1}^{365} \sum_{j=1}^{m(n)} \left\{ \log \left[\frac{\alpha(n)}{\beta(n)} \exp[-u_{nj}/\beta(n)] \cdot \left(\frac{1-\alpha(n)}{\delta(n)}\right) \exp(-u_{nj}/\delta(n))\right]\right\} \tag{11}$$

where $m(n)$ = number of wet days n; $n = 1, 2 \cdots 365$ for the period of record, u_{nj} = the transformed precipitation for the jth wet day for day n.

Each parameter in (9)–(11) is expressed in the finite Fourier series form as specified by (4). Let $\boldsymbol{\theta}_d$, $\boldsymbol{\theta}_w$, and $\boldsymbol{\theta}_M$ be vectors whose elements are the coefficients of the Fourier series describing the parameter set for the dry-dry transitions, the wet-dry transitions, and the mixed exponential distributions, respectively. The objective is to find the estimate $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ that maximizes log L in (9)–(11).

The likelihood ratio test [cf. *Hoel*, 1971; *Mielke and Johnson*, 1973] was used to determine if an added harmonic was significant. For example, to determine if the first harmonic is significant in describing the seasonal variation in $p_{00}(k)$, we test the null hypothesis

$$H_0: \quad \boldsymbol{\theta}_d = \boldsymbol{\theta}_d' = (\gamma_{10}) \tag{12}$$

against the alternative

$$H_1: \quad \boldsymbol{\theta}_d = (\gamma_{10}, C_{11}, \phi_{11}) \tag{13}$$

Let $L(\mathbf{x}, \boldsymbol{\theta}_d')$ be the maximum likelihood function when H_0 is true and $L(\mathbf{x}, \boldsymbol{\theta}_d)$ be the maximum likelihood function under the alternative hypothesis. Under certain regularity conditions, the statistic, $-2 \log_e \{L(\mathbf{x}, \boldsymbol{\theta}_d')/L(\mathbf{x}, \boldsymbol{\theta}_d)\}$ has a distribution that approaches the chi square distribution with 2 degrees of freedom for large sample size. We accepted the null hypothesis if the probability of obtaining a greater test statistic was smaller than 0.05. As *Woolhiser and Pegram* [1979] have noted, the true level of significance is somewhat different because of repeated testing and problems with dependence between parameters.

It can be shown that the amplitudes and phase angles of the significant harmonics for each parameter are not independent. Furthermore, the parameters α, β, and δ are also not independent. For this reason, the order in which the parameters are analyzed will affect which harmonics are declared significant by the likelihood ratio test and also the final log-likelihood value. Because one goal of this research is to present a regionalized precipitation model, it is important that consistent procedures be followed. Accordingly, three stations were analyzed in detail using different optimization sequences. The sequence leading to the highest final log-likelihood value was selected. The order in which parameters were included was δ, β, and α for both harmonics. The order for the Markov Chain parameters is not important, because they are independent. The first through the fourth harmonics were included for each parameter in that order.

The means of the Markov chain parameters (γ_{10}, γ_{20}) need not be optimized simultaneously, because they are independent; however, they are not independent of their associated amplitudes and phase angles. Therefore a three-parameter (mean, amplitude, and phase angle) optimization was made when the first harmonic was studied. If the first harmonic was not significant, the mean was studied along with the second harmonic. If the first harmonic was significant, the mean was fixed, and subsequent optimizations included only the amplitude and phase angle of higher harmonics. No improvement was obtained by optimizing the mean along with all harmonics, but the computer time was increased.

Estimation of the three mixed exponential distribution parameters γ_{30}, γ_{40}, γ_{50} is more difficult because the three mean values are not independent. Likewise, the mean value of each parameter is not independent of its associated amplitudes and phase angles. Thus a large number of procedures can be followed in the optimization process. *Woolhiser and Roldán* [1982] demonstrated that the best procedure is to optimize the three means simultaneously and then to sequentially optimize the amplitude and phase angle of each harmonic. We found that the procedure described by *Woolhiser and Roldán* [1982] led to biased estimates of the expected annual precipitation as calculated by (5) but that this bias could be nearly eliminated by a second round of optimization retaining all parameters previously declared significant.

SPATIAL CHARACTERISTICS OF THE LOCAL DAILY PRECIPITATION PROCESS

The spatial (geographical) character of the local daily precipitation process can be described by the random fields $\gamma_{i0}(\mathbf{u})$, $c_{ij}(\mathbf{u})$, and $\phi_{ij}(\mathbf{u})$, where $\mathbf{u}$ is the vector of spatial coordinates $\mathbf{u} = (x, y)$, $i = 1, 2 \cdots 5$, and $j = 1, 2 \cdots m_i$. Because precipitation measurements are made at points in space, we must infer the properties of the continuous random fields from the parameter estimates at M stations with coordinates $\mathbf{u}_m = (x_m, y_m)$ ($m = 1, 2, \cdots M$). It should be emphasized here that we are considering only the spatial variability of the local daily precipitation process, X_t Y_t, as evidenced by the spatial character of the fields describing the seasonal variation of parameters of this process. The more difficult time-space daily precipitation process is beyond the scope of this investigation.

Creutin and Obled [1982] presented a comprehensive review and evaluation of mapping techniques for rainfall fields. The parameter fields under discussion here must be viewed in the context of a single realization, and therefore the mapping techniques must be selected from those that they classified as spatial methods rather than climatological methods. Included in this category are

1. The nearest neighbor method: The estimated value of any given point is taken as the observed value at the nearest neighboring station.
2. The arithmetic mean: It is assumed that the parameter is constant over a particular region and can be estimated by the average of the observed values within the region.
3. Spline-surface fitting: This consists of finding the surface interpolating the observed points, which also satisfies an optimal "smoothness" criterion. If only three observed points are considered, this method reduces to a linear interpolation.
4. Kriging method: The value at an ungaged point is estimated as a linear combination of n surrounding observed values, which minimizes the estimation variance.

Each of these methods may be used to estimate point values. The point values estimated by methods 3 and 4 could be used to prepare maps of parameter isolines to provide a regional description of the daily local precipitation process. As we have pointed out previously, some of the parameter fields under consideration are not independent; however, as a first approximation, we will neglect this dependence in the following analyses.

SPATIAL CHARACTERISTICS OF PARAMETERS: SOUTH DAKOTA

Precipitation data for the state of South Dakota were used to examine the spatial variability of the Fourier coefficients describing the MCME parameters. This region is ideal for such a study because there is a substantial gradient of precipitation from southeast to northwest and orographic effects should be minor, except in the Black Hills region in the southwestern corner of the state.

Figure 1 is a map of the state of South Dakota showing the 16 stations used for parameter estimation and the four sta-

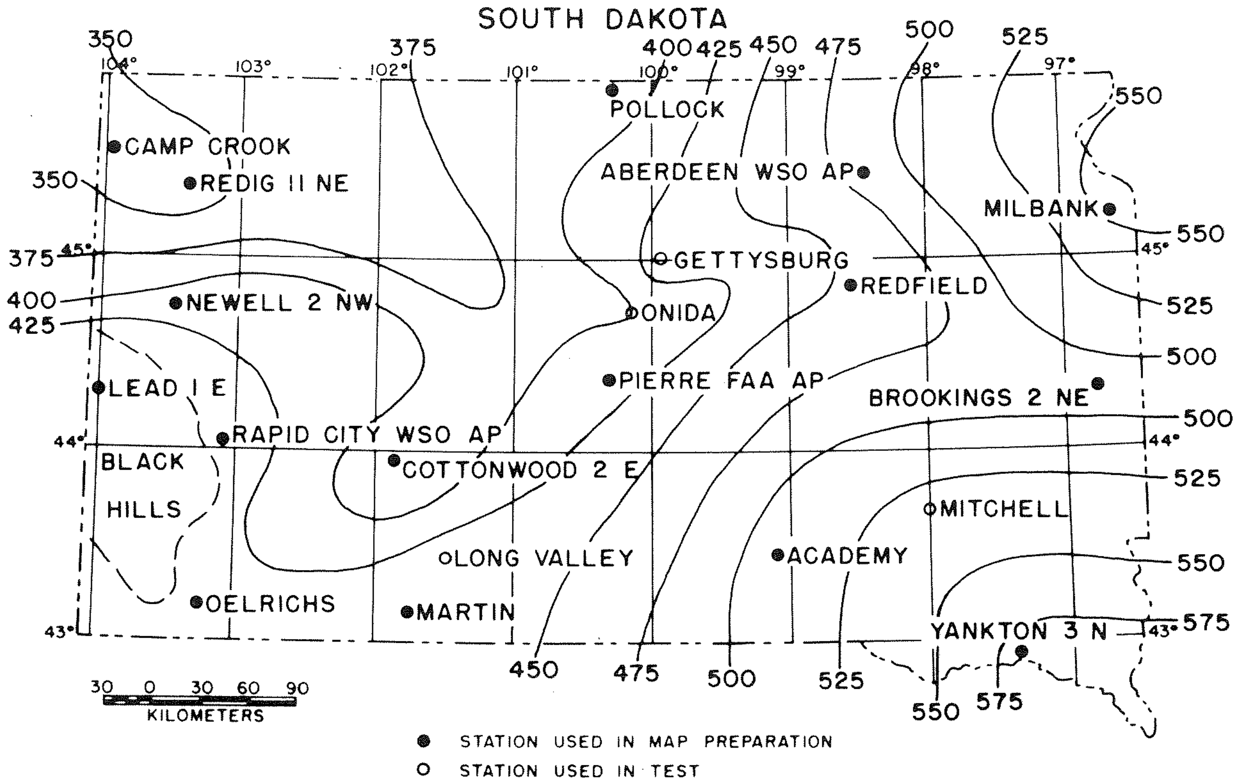

Fig. 1. Location map and average annual precipitation.

tions used for testing alternative interpolation methods. Isohyets of mean annual precipitation calculated from National Weather Service daily data are also shown. Except for Martin, Gettysburg, and Long Valley, these means are based on a 40-year period of record beginning March 1, 1928, and ending February 28, 1959. The observation time, total number of wet days, mean annual number of wet days, mean precipitation on wet days, and elevation above sea level for all stations are shown in Table 1. The data from Aberdeen and Rapid City are from recording gages; all other stations had 8-inch (20.3 cm) standard rain gages read at 24-hour intervals. The means, amplitudes, and phase angles for the Markov chain parame-

TABLE 1. Meteorological Information for Stations Analyzed (1928–1959 Except as Noted).

Station	Observation Time, LT	Total Number of Wet Days	Mean Precipitation on Wet Days, mm	Mean Annual Number of Wet Days	Elevation, m	Mean Annual Precipitation, mm
		Base Stations				
Aberdeen	0000	3270	6.20	82	395	507
Academy	1800	2324	8.74	58	511	508
Brookings	0700	3137	6.27	78	500	492
Camp Crook	1700	2439	5.64	61	951	344
Cottonwood	1700	2592	5.99	65	736	388
Lead 1 E	1800	4602	5.33	115	1916	613
Martin (1934–1973)	1700	2559	6.22	69	1089	430
Milbank	0700	3027	7.34	76	349	555
Newell	0800	3492	4.65	87	875	406
Oelrichs	1800	2009	8.53	50	1017	428
Pierre	0000	3282	5.11	82	529	419
Pollock	1900	2200	7.04	55	498	387
Rapid City	0000	3788	4.52	95	965	428
Redfield	1800	2702	6.76	68	395	457
Redig	SS	2881	4.78	72	936	344
Yankton	0700	3256	7.11	81	387	579
		Test Stations				
Gettysburg (1931–1970)	1800	2696	6.36	67	634	428
Long Valley (1927–1966)	1800	2559	7.52	69	753	481
Mitchell 2 SW	0700	3341	6.45	84		539
Onida 4 NW	1800	2132	7.52	53	564	400

SS, observation time near sunset.

TABLE 2. Fourier Coefficients for the Markov Chain Parameter p_{00}

Station	Mean γ_{10}	C_{11}	ϕ_{11}, rad	C_{12}	ϕ_{12}, rad	C_{13}	ϕ_{13}, rad	C_{14}	ϕ_{14}, rad
				Base Stations					
Aberdeen	0.8166	0.0578	2.9906	0.0171	0.3606	*	*	*	*
Academy	0.8645	0.0878	2.7382	0.0146	1.4611	*	*	*	*
Brookings	0.8191	0.0763	2.7852	0.0156	0.9210	*	*	*	*
Camp Crook	0.8652	0.0597	2.9728	0.0145	0.8075	0.0254	0.7695	0.0126	−2.0938
Cottonwood	0.8554	0.0752	2.8645	0.0191	1.2622	0.0120	−0.0408	*	*
Lead 1 E	0.7502	0.0626	4.0111	0.0229	1.1149	0.0147	−2.1001	*	*
Martin	0.8425	0.0732	2.9288	0.0190	0.6847	*	*	*	*
Milbank	0.8299	0.0769	2.6661	0.0165	0.7123	*	*	*	*
Newell	0.8107	0.0684	3.0760	0.0168	1.4434	*	*	*	*
Oelrichs	0.8849	0.0663	2.8537	0.0144	2.0121	0.0111	−1.8125	*	*
Pierre	0.8202	0.0681	3.0101	0.0239	0.5885	0.0124	−0.5478	*	*
Pollock	0.8710	0.0599	2.7472	0.0195	0.8586	0.0116	−0.7555	*	*
Rapid City	0.7973	0.0837	3.1115	0.0237	0.8546	*	*	*	*
Redfield	0.8453	0.0785	2.6493	0.0181	0.7983	*	*	*	*
Redig	0.8404	0.0695	2.8531	0.0120	1.0363	0.0124	−0.7198	0.0110	−2.3941
Yankton	0.8185	0.0917	2.7340	*	*	*	*	0.0189	−1.441
				Test Stations					
Gettysburg	0.8463	0.0724	2.8390	0.0187	0.9320	*	*	0.0134	−2.2531
Long Valley	0.8700	0.0627	3.1709	*	*	*	*	*	*
Mitchell	0.8199	0.0835	2.7721	*	*	*	*	0.0126	−1.5337
Onida 4	0.8778	0.0604	2.7861	0.0117	1.2786	0.0095	−0.1680	*	*

*Harmonic not significant at 0.05 level.

ters and the mixed exponential distribution parameters are presented in Tables 2 through 5.

A preliminary analysis revealed that for most stations the parameter α was constant throughout the year. Therefore no higher harmonics were considered in all optimizations. This procedure prevented potentially severe interactions between harmonics in α and the other ME parameters.

To provide a visual impression of the spatial characteristics of the Fourier coefficients, isopleth maps of the means, γ_{i0}, $i = 1, 2, \cdots 5$, were prepared by drawing smooth curves through points obtained by linear interpolation between the coefficient values for adjacent stations. Parameters estimated for the test stations were not used in constructing the isopleths. Isopleth maps of the means of p_{00} and p_{10} for the Markov chain are shown in Figures 2 and 3. It is apparent from these figures that there are significant spatial variations in these occurrence process parameters. However, it must be emphasized that each optimized coefficient includes a sampling error term due to the finite length of record and also includes a measurement error term. The annual mean probability of a wet day following a dry day $(1 - \bar{p}_{00})$ is greatest in the southeast and in the Black Hills region and is lowest in the northwest and southwest. This is generally true of the mean wet-wet transition probability $(1 - \bar{p}_{10})$, except for the relative maximum at Pierre, in the center of the state. The trends from the northwest to the southeast are probably real; however,

TABLE 3. Fourier Coefficients for the Markov Chain Parameter p_{10}

Station	Mean γ_{20}	C_{21}	ϕ_{21}, rad	C_{22}	ϕ_{22}, rad	C_{23}	ϕ_{23}, rad	C_{24}	ϕ_{24}, rad
				Base Stations					
Aberdeen	0.6332	0.0494	−2.8574	0.0398	2.2510	*	*	*	*
Academy	0.7143	0.0978	−3.0031	0.0815	2.8908	*	*	*	*
Brookings	0.6613	0.0840	−3.1717	0.0653	2.8360	*	*	*	*
Camp Crook	0.6619	0.0769	−3.0811	0.0603	2.2251	*	*	*	*
Cottonwood	0.6843	0.1302	−3.1392	0.0469	2.6142	*	*	*	*
Lead 1 E	0.5449	0.0809	−2.6774	0.0391	2.1198	*	*	*	*
Martin	0.6625	0.0860	−2.9634	0.0711	2.4512	0.0535	0.0527	*	*
Milbank	0.6444	0.0556	−3.4206	0.0605	2.4631	*	*	*	*
Newell	0.6035	0.0794	−3.3779	0.0653	2.1368	0.0542	0.1564	0.0327	2.5102
Oelrichs	0.7184	0.0688	−2.7602	0.0520	2.4645	*	*	*	*
Pierre	0.6170	0.0586	−2.9438	0.0466	1.9987	*	*	*	*
Pollock	0.7308	0.0782	−3.3887	0.0412	2.6174	0.0414	−0.0697	*	*
Rapid City	0.5786	0.0905	−2.9413	0.0458	2.3506	0.0316	−0.5810	*	*
Redfield	0.6755	0.0733	−2.9872	0.0865	2.3667	*	*	*	*
Redig	0.6548	0.0960	−3.5296	0.0553	2.1858	0.0355	−0.0093	*	*
Yankton	0.6282	0.0824	−3.0682	0.0465	2.6165	*	*	*	*
				Test Stations					
Gettysburg	0.6831	0.0974	−3.1735	0.0639	2.1576	*	*	*	*

TABLE 4. Fourier Coefficients for the Mixed Exponential Distribution Parameters α and β

Station	α γ_{30}	β γ_{40}, mm	C_{41}, mm	ϕ_{41}, rad	C_{42}, mm	ϕ_{42}, rad	C_{43}, mm	ϕ_{43}, rad	C_{44}, mm	ϕ_{44}, rad
Base Stations										
Aberdeen	0.3873	1.084	0.652	−0.3361	*	*	*	*	*	*
Academy	0.6034	5.158	0.844	−0.6221	*	*	*	*	*	*
Brookings	0.4061	1.190	0.805	−0.9419	*	*	*	*	*	*
Camp Crook	0.5135	1.998	1.116	−0.5970	*	*	*	*	*	*
Cottonwood	0.5135	1.752	0.378	−0.4025	*	*	*	*	*	*
Lead 1 E	0.5777	1.985	0.031	−0.5084	0.2797	−1.2888	*	*	*	*
Martin	0.6233	2.378	0.639	−0.5896	*	*	*	*	*	*
Milbank	0.4482	2.323	*	*	*	*	*	*	*	*
Newell	0.4568	0.916	0.478	−0.7552	0.1311	−1.4224	*	*	*	*
Oelrichs	0.8484	5.965	1.830	−0.4766	0.8197	−0.9568	0.7493	−2.606	*	*
Pierre	0.4462	0.816	0.465	−0.7626	0.0996	−0.8102	*	*	*	*
Pollock	0.6290	3.964	1.511	−0.3202	*	*	*	*	*	*
Rapid City	0.4197	0.696	0.408	−0.5410	0.0950	−0.6382	0.1072	−2.942	*	*
Redfield	0.5332	2.199	0.986	−0.4850	*	*	0.3922	−2.858	*	*
Redig	0.5027	1.610	*	*	*	*	*	*	*	*
Yankton	0.4988	1.364	0.573	−0.6232	*	*	*	*	*	*
Test Stations										
Gettysburg	0.5629	2.539	0.835	−0.3344	*	*	*	*	*	*
Long Valley	0.6589	4.553	1.426	−0.4382	*	*	*	*	*	*
Mitchell	0.3857	0.934	3.597	−0.5935	*	*	*	*	*	*
Onida 4	0.6240	4.605	1.245	−0.3849	*	*	*	*	*	*

*Not significant at 0.05 level.

some of the differences may be due to the observation time or to the possibility that some cooperative observers are reporting too few wet days.

Isopleth maps for the means of α, β, and δ for the mixed exponential distribution are shown in Figures 4, 5, and 6. The parameter α shows significant spatial variation with a range from 0.4 to 0.8. Both β and δ show significant spatial variations as well. There is some similarity between the patterns of the isopleths for α and the mean values of β and δ. Because of their dependence, this is to be expected. Concentrations of isopleths are apparent near Oelrichs and Academy, suggesting either rapid changes in the distribution of amounts or significant measurement errors.

COMPARISON OF SELECTED TECHNIQUES FOR ESTIMATING PARAMETERS AT UNGAGED LOCATIONS

Fourier coefficients for each of the five parameters in the MCME model were estimated for the four test stations shown in Table 1. The following techniques were used: (1) nearest

TABLE 5. Fourier Coefficients for the Mixed Exponential Distribution Parameter δ

Station	γ_{50}, mm	C_{51}, mm	ϕ_{51}, rad	C_{52}, mm	ϕ_{52}, rad	C_{53}, mm	ϕ_{53}, rad	C_{54}, mm	ϕ_{54}, rad
Base Stations									
Aberdeen	8.915	5.161	−0.8031	*	*	*	*	*	*
Academy	11.814	7.008	−0.8322	*	*	*	*	*	*
Brookings	8.963	5.878	−0.8724	*	*	*	*	*	*
Camp Crook	8.307	5.385	−0.7437	*	*	*	*	*	*
Cottonwood	8.830	3.797	−0.7052	*	*	*	*	*	*
Lead 1 E	9.792	6.358	−0.7911	0.6553	−1.8196	0.6248	−3.0335	*	*
Martin	11.361	10.777	−0.6848	*	*	*	*	*	*
Milbank	10.279	4.732	−0.7504	*	*	*	*	*	*
Newell	7.084	4.737	−0.7033	*	*	0.5893	−2.7910	*	*
Oelrichs	16.187	14.636	−0.8286	*	*	*	*	*	*
Pierre	7.716	4.206	−0.6764	0.6553	0.1056	*	*	*	*
Pollock	9.901	8.423	−0.8750	*	*	1.5443	−3.0897	*	*
Rapid City	6.741	4.463	−0.6862	0.4191	−0.7159	0.5105	−2.868	*	*
Redfield	9.797	4.097	−0.7270	*	*	*	*	*	*
Redig	6.292	3.620	−0.6661	*	*	*	*	*	*
Yankton	10.958	5.027	−0.8480	*	*	*	*	*	*
Test Stations									
Gettysburg	9.705	6.322	−0.6921	*	*	*	*	*	*
Long Valley	11.059	10.112	−0.6451	*	*	*	*	*	*

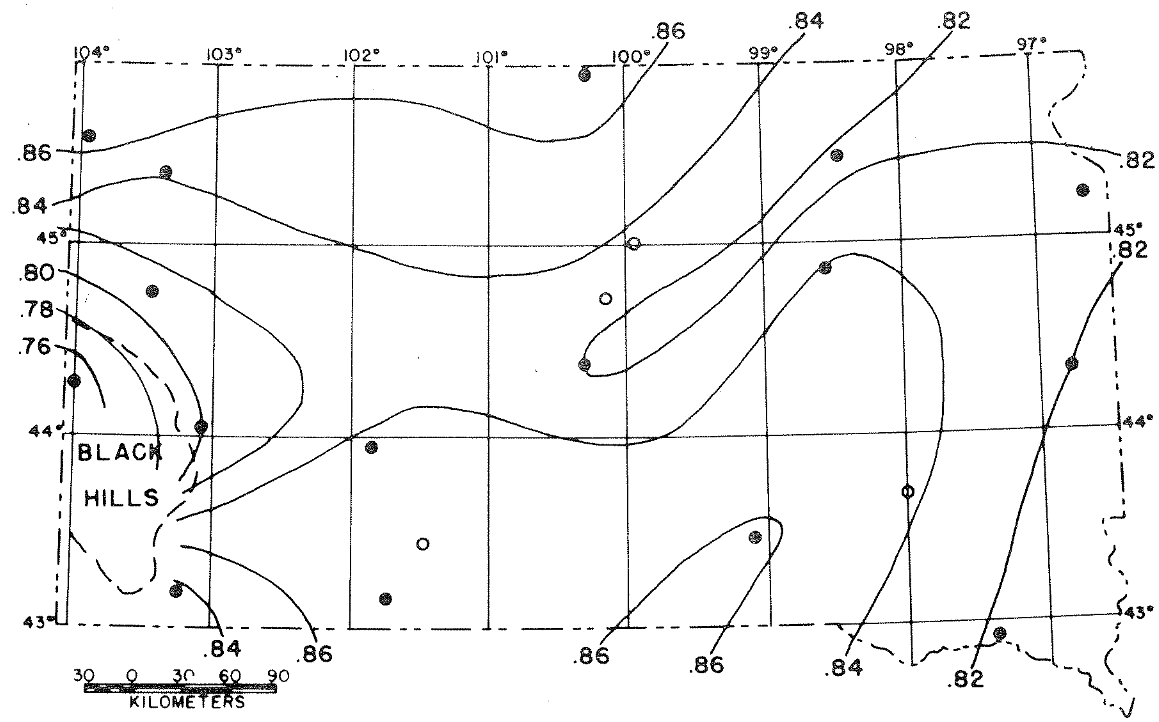

Fig. 2. Isopleth map of annual mean p_{00} (γ_{10}).

neighbor, (2) the arithmetic mean of the six nearest stations, (3) a spline-surface fit to the six nearest stations using the technique described by *Creutin and Obled* [1982] and attributed to *Duchon* [1976] and *Paihua and Utreras* [1978], and (4) a linear interpolation using the three nearest stations defining a triangle that includes the station.

The estimated coefficients, using each of the above methods, are shown for the Markov chain and the mixed exponential distribution in Tables 6 and 7, respectively.

The techniques were compared by calculating the log-likelihood functions using the estimated coefficients and (9)–(11) with precipitation data for each test station. The maximum likelihood (ML) functions for each test station can be utilized to test the following hypothesis:

$$H_0: \quad \omega_1 = (\boldsymbol{\theta}_1{}') \qquad \omega_2 = (\boldsymbol{\theta}_2{}')$$

against the alternative

$$H_1: \quad \omega_1 = (\boldsymbol{\theta}_1) \qquad \omega_2 = (\boldsymbol{\theta}_2)$$

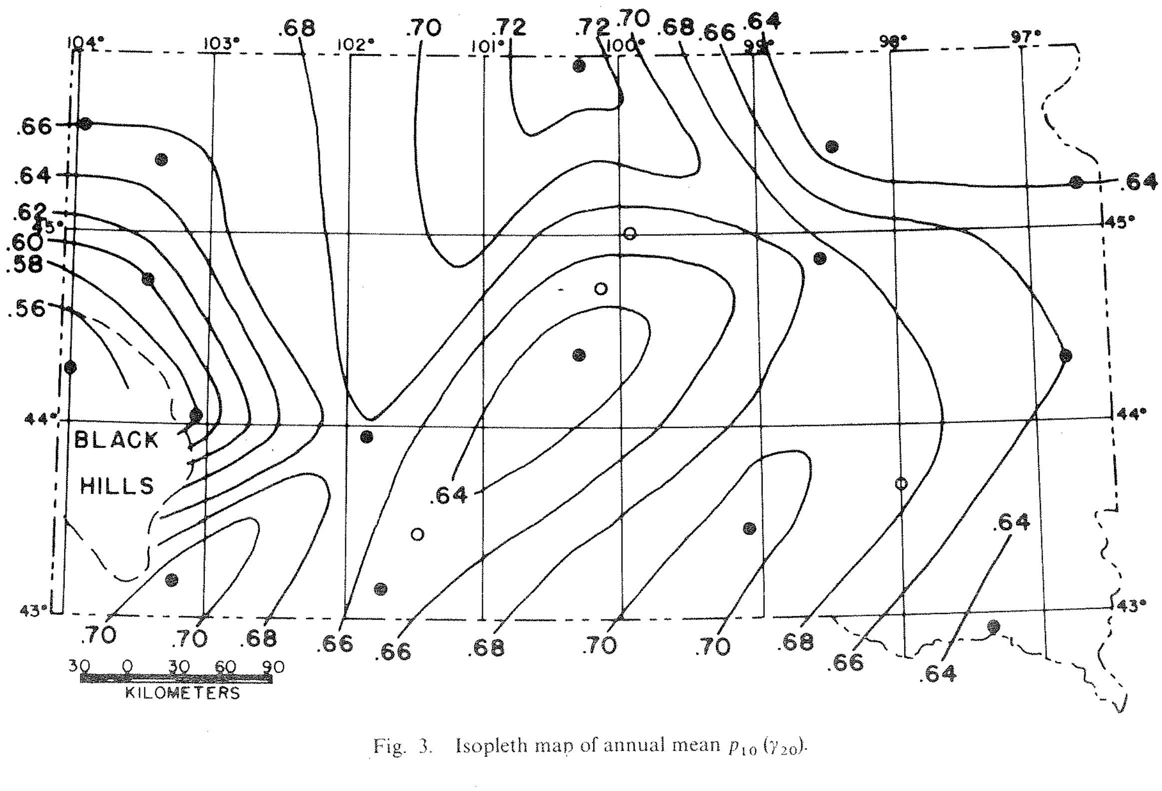

Fig. 3. Isopleth map of annual mean p_{10} (γ_{20}).

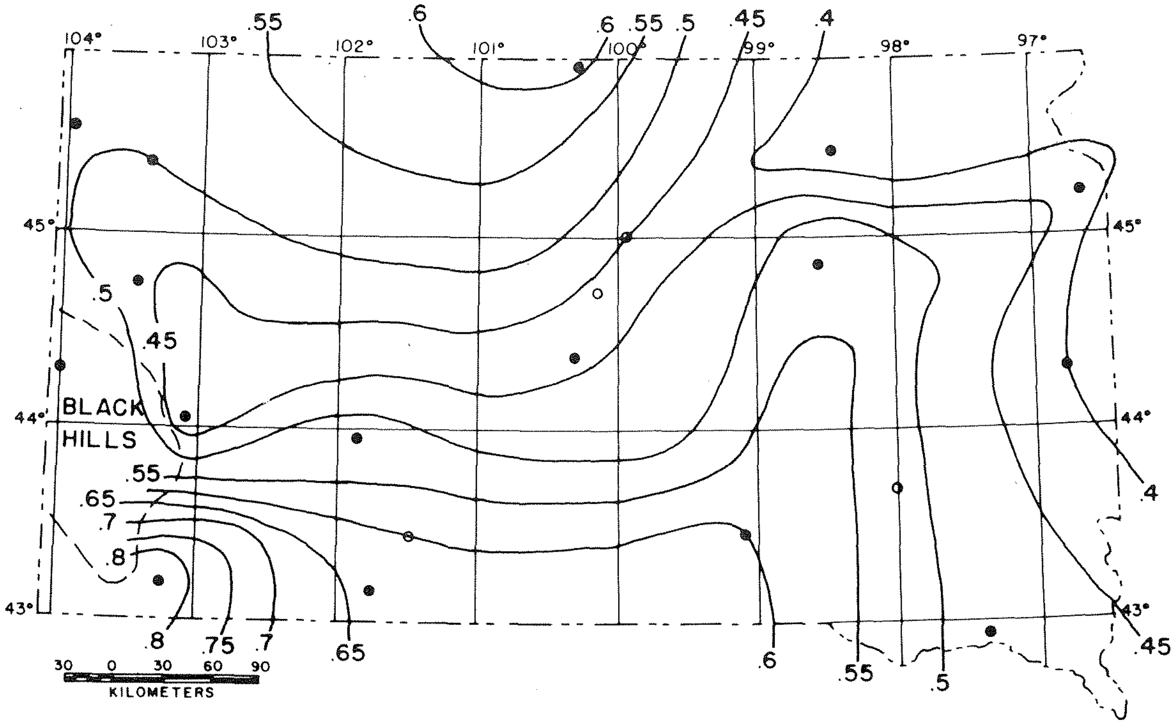

Fig. 4. Isopleth map of mean α (γ_{30}).

where the subscripts 1 and 2 refer to the occurrence process and ME distribution, respectively, and $\boldsymbol{\theta}_1'$ and $\boldsymbol{\theta}_2'$ refer to parameters estimated by one of the four techniques discussed previously; $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are the parameter vectors estimated from the actual record at the test station by ML techniques with the constraint that each parameter of the MCME model will have no more harmonics than the neighboring stations (i.e., both p_{00} and p_{10} are allowed two harmonics; both β and δ are allowed one harmonic, and α is described only by the mean value).

The likelihood ratio statistic

$$\lambda_i = -2 \log_e \{L(\mathbf{x}, \boldsymbol{\theta}_i')/L(\mathbf{x}, \boldsymbol{\theta}_i)\} \qquad i = 1, 2$$

is approximately chi square with 10 degrees of freedom for the Markov chain and 7 degrees of freedom for the mixed exponential distribution.

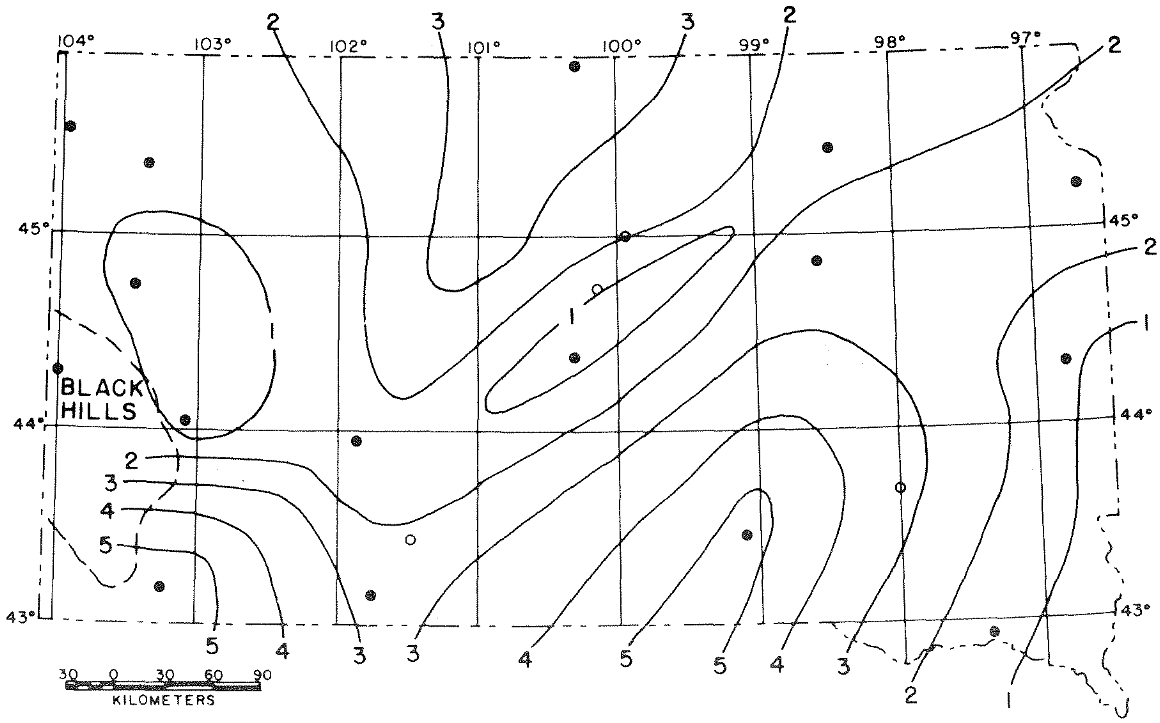

Fig. 5. Isopleth map of mean β (γ_{40}), mm.

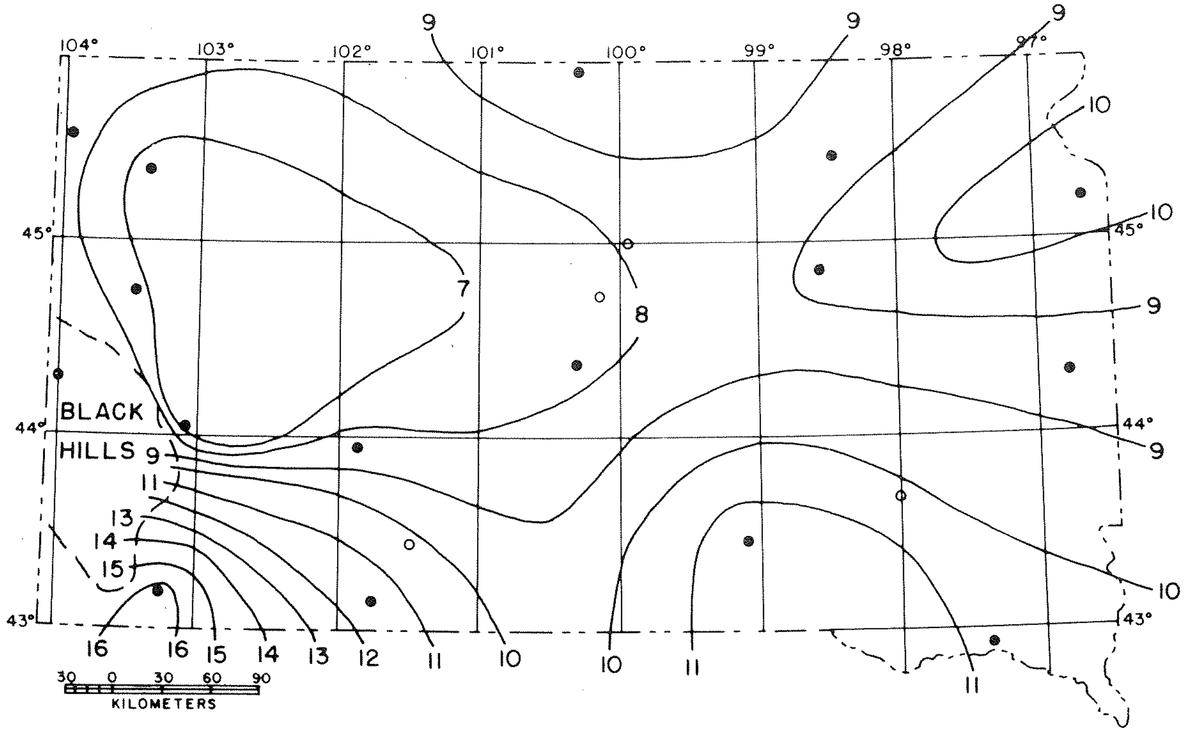

Fig. 6. Isopleth map of mean δ (γ_{50}), mm.

TABLE 6. Fourier Coefficients for Test Stations: Markov Chain

Estimation Technique	P_{00} Mean γ_{10}	P_{00} C_{11}	P_{00} ϕ_{11}, rad	P_{00} C_{12}	P_{00} ϕ_{12}, rad	P_{10} Mean γ_{20}	P_{10} C_{21}	P_{10} ϕ_{21}, rad	P_{10} C_{22}	P_{10} ϕ_{22}, rad	P_{10} C_{23}	P_{10} ϕ_{23}, rad	Log Likelihood
						Long Valley Station							
1 (Martin)	0.8425	0.0732	2.9288	0.0190	0.6847	0.6625	0.0860	−2.9640	0.0711	2.4512	0.0549	0.0527	−6014.54
2	0.8441	0.0757	2.9178	0.0191	1.1439	0.6625	0.0886	−2.9586	0.0573	2.4617			−6009.55
3	0.8449	0.0742	2.9042	0.0193	0.7787	0.6682	0.1013	−3.0405	0.0649	2.5028			−6008.41
4	0.8390	0.0754	2.9184	0.0200	0.6443	0.6598	0.1072	−3.0758	0.0653	2.5138			−6029.44
ML	0.8700	0.0627	3.1710	*	*	0.7123	0.1018	−3.0391	0.0635	2.2982			−5952.78
						Mitchell Station							
1 (Academy)	0.8645	0.0878	2.7382	0.0146	1.4611	0.7143	0.0978	−3.0031	0.0815	2.8908	*	*	−7479.54
2	0.8329	0.0799	2.7629	0.0158	1.0407	0.6568	0.0753	−3.0991	0.0645	2.5286			−7304.78
3	0.8353	0.0821	2.8404	0.0151	1.2818	0.6669	0.0878	−2.9932	0.0627	2.6796			−7315.81
4	0.8432	0.0845	2.7528	0.0137	1.3295	0.6847	0.0910	−3.0796	0.0713	2.8336			−7347.85
ML	0.8198	0.0835	2.7721	*	*	0.6154	0.0630	−2.9380	0.0446	2.5872			−7283.02
ML 4th harmonic				0.0126	−1.5336								
						Gettysburg Station							
1 (Redfield)	0.8453	0.0785	2.6493	0.0181	0.7983	0.6755	0.0733	−2.9872	0.0865	2.3667	*	*	−6575.26
2	0.8413	0.0715	2.8003	0.0183	0.7966	0.6692	0.0688	−3.0643	0.0594	2.4313			−6572.04
3	0.8372	0.0678	2.8345	0.0206	0.6340	0.6593	0.0653	−2.9579	0.0595	2.2244			−6579.63
4	0.8053	0.0546	3.1664	0.0211	0.2978	0.6024	0.0438	−2.8508	0.0217	2.0336			−6716.52
ML	0.8463	0.0724	2.8390	0.0187	0.9320	0.6831	0.0974	−3.1735	0.0639	2.1576	*	*	−6561.01
ML 4th harmonic				0.0134	−2.2531								
						Onida Station							
1 (Pierre)	0.8202	0.0681	3.0101	0.0239†	0.5885†	0.6170	0.0586	−2.9438	0.0466	1.9987	*	*	−5929.50
2	0.8455	0.0712	2.8333	0.0187	0.8882	0.6758	0.0812	−3.0174	0.0571	2.4565			−5778.14
3	0.8190	0.0626	3.0476	0.0234	0.4892	0.6192	0.0517	−2.9363	0.0372	2.0130			−5944.24
4	0.8328	0.0672	2.9270	0.0225	0.6612	0.6452	0.0638	−2.9951	0.0486	2.1564			−5828.23
ML	0.8778	0.0604	2.7861	0.0117	1.2786	0.7251	0.1048	−3.2027	0.0579	1.8452	*	*	−5703.85
ML 3rd harmonic				0.0095	−0.1680								

1, Nearest neighbor; 2, Arithmetic mean; 3, Spline; 4, Linear interpolation; ML, Maximum likelihood.
*Harmonic not significant at 0.05 level.
†Third harmonic for Pierre; AMP, 0.01238; PHS, −0.54779.

TABLE 7. Fourier Coefficients for Test Stations: Mixed Exponential

Estimation Technique	α: γ_{30}	β: γ_{40}, mm	β: C_{41}, mm	β: ϕ_{41}, rad	β: C_{42}, mm	β: ϕ_{42}, rad	β: C_{43}, mm	β: ϕ_{43}, rad	δ: γ_{50}, mm	δ: C_{51}, mm	δ: ϕ_{51}, rad	δ: C_{52}, mm	δ: ϕ_{52}, rad	Log Likelihood
						Long Valley Station								
1 (Martin)	0.6233	2.3779	0.6388	−0.5896	*	*	*	*	11.3622	10.7775	−0.6848	*	*	591.93
2	0.5757	2.7940	0.7620	−0.5657	*	*	*	*	10.4419	7.4828	−0.7356	*	*	724.18
3	0.5689	1.9533	0.4623	−0.5457	*	*	*	*	10.1752	8.0264	−0.6833	*	*	609.07
4	0.5332	1.3945	0.2896	−0.5376	*	*	*	*	9.3599	7.1933	−0.6635	*	*	512.32
ML	0.6589	4.5527	1.4257	−0.4382	*	*	*	*	11.0589	10.3010	−0.8079	*	*	772.05
						Mitchell Station								
1 (Academy)	0.6035	5.1580	0.8443	−0.6221	*	*	*	*	11.8140	7.0089	−0.8322	*	*	1485.68
2	0.4893	2.1742	0.6681	−0.7007	*	*	*	*	9.9212	5.1613	−0.7844	*	*	1826.08
3	0.5059	2.5883	0.7417	−0.7462	*	*	*	*	10.0254	6.0274	−0.8263	*	*	1779.50
4	0.5244	3.3274	0.7925	−0.7260	*	*	*	*	10.7696	6.3576	−0.8475	*	*	1686.01
ML	0.3857	0.9337	0.3597	−0.5935	*	*	*	*	9.0279	4.8913	−0.7555	*	*	1921.90
						Gettysburg Station								
1 (Redfield)	0.5332	2.1994	0.9865	−0.4850	*	*	0.3922	−2.8577	9.7973	4.0970	−0.7270	*	*	1390.79
2	0.5079	2.5908	0.7976	−0.5492	*	*	*	*	9.7384	5.6058	−0.7773	*	*	1401.04
3	0.5023	1.6764	0.9169	−0.4999	*	*	*	*	8.7808	4.8641	−0.7389	*	*	1381.16
4	0.3553	0.3785	0.3708	−0.5392	*	*	*	*	7.7216	5.0571	−0.7581	*	*	1069.43
ML	0.5629	2.5387	0.8354	−0.3344	*	*	*	*	9.7051	6.3218	−0.6922	*	*	1407.48
						Onida Station								
1 (Pierre)	0.4462	0.8156	0.4651	−0.7626	0.0996	−0.81020	*	*	7.7158	4.2075	−0.6764	0.6563	0.1056	390.46
2	0.5188	2.4943	0.8052	−0.4881	*	*	*	*	9.4945	5.4483	−0.7698	*	*	652.30
3	0.4374	0.9246	0.5486	−0.6965	*	*	*	*	7.7470	4.9530	−0.7140	*	*	440.53
4	0.4911	1.5799	0.7247	−0.6488	*	*	*	*	8.3287	5.0800	−0.7217	*	*	560.76
ML	0.6240	4.6050	1.2454	−0.3840	*	*	*	*	10.8791	10.1117	−0.6450	*	*	689.25

1, Nearest neighbor; 2, Arithmetic mean; 3, Spline; 4, Linear interpolation; ML, Maximum likelihood.

For the occurrence process, method 2 (arithmetic mean of six nearest stations) gave the highest likelihood function for three of the four stations and was second best for the fourth. However, the null hypothesis was rejected ($p = 0.05$) for all techniques for all stations, which means that all estimation techniques gave Fourier coefficient values that were statistically different from those estimated using the real data at the test station. The arithmetic mean provided the best estimators for the distribution of amounts for all stations, but the null hypothesis was rejected for all cases except for the arithmetic mean estimates at Gettysburg.

Although the number of stations is marginal [c.f. *Hughes and Lettenmaier*, 1981], the regionalization technique of kriging was also examined. Semivariograms were computed for the means of the parameters p_{00} and p_{10}, using the universal kriging program described by *Skrivan and Karlinger* [1979], and are shown in Figure 7. The variograms are flat for both parameters, suggesting a significant "nugget" effect. This rather large nugget variance may be caused by a number of factors, including real mesoscale differences in the precipitation regime at a scale much smaller than the spacing of the data points, measurement errors, observer bias, time of reading the gages, and modeling errors. A normalizing transformation of the form

$$r_i = \log\left(\frac{p_{i0}}{1 - p_{i0}}\right) \qquad i = 0, 1$$

was also tried, but the shape of the empirical semivariogram did not change. The rather poor performance of all interpolation techniques indicates potentially serious problems. Therefore the factors which may have contributed were examined in more detail.

Parameter Identifiability

Because all Fourier coefficients are not considered simultaneously in the optimization process and the coefficients are, in fact, dependent, it is possible that some of the spatial variability in individual coefficients is due to convergence to a local optimum. Parameter sampling variability is always present as well, although it will decrease as record length increases. An empirical examination of parameter identifiability was per-

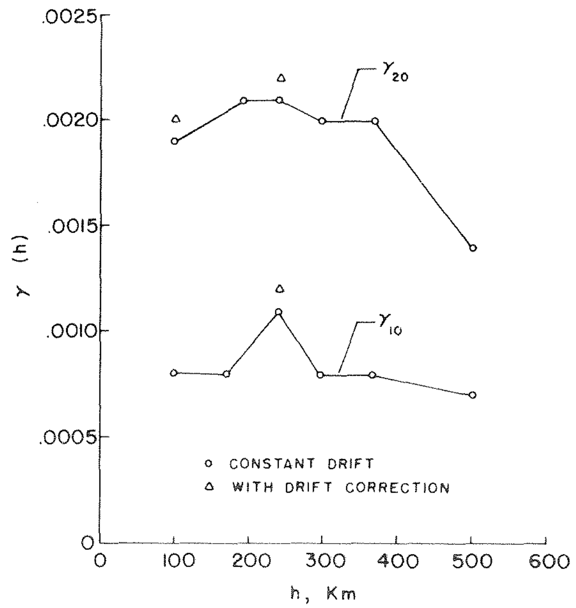

Fig. 7. Semivariograms for Markov chain parameters γ_{10} and γ_{20}.

formed by simulating 10 sets of 40-year records using Fourier coefficients identified for Pierre. Fourier coefficients were then estimated for each simulated record. The means and standard deviations of the Fourier cofficients are compared with the theoretical coefficients in Table 8.

The statistics presented in Table 8 show that there is little sampling variability in the Fourier coefficients for the Markov chain parameters p_{00} and p_{10} and for the mean values of α, β, and δ. However, there is a significant problem in identifying the Fourier coefficients for the first and second harmonics for β and δ. Although two harmonics were present for δ in the simulation model, the first harmonic was identified as significant only once, and the second was not identified as significant for any of the 10 samples. This demonstrates the strong dependence between the parameters α, β, and δ and indicates that it is possible to reach local optima. Thus we conclude that the noise in the Markov chain parameter fields and in the means of α, β, and δ is probably not due to sampling variation or problems of parameter identification but that parameter identification problems could be significant for the second and higher harmonics for β and all harmonics for δ for the ME parameter fields.

Effects of Observation Time

From Table 1, we see that the observation time is not the same for each station. If there is a substantial diurnal variation in the rainfall process, differences in the observation times could cause differences in both the number of wet days and the distribution of precipitation amounts per day. To investigate this possibility, we obtained hourly precipitation data for Rapid City and Aberdeen from the National Climatic Data Center, NOAA.

The frequency of precipitation occurrence during each hour was estimated for each 14-day period and for the year (see Figure 8). Three daily records, beginning at midnight, 0700, and 1800, were assembled from the hourly records for each station, and the Fourier coefficients were estimated for each record. Likelihood ratio tests showed that the null hypothesis could not be rejected at the 5% level for the ME model for Aberdeen; however, the null hypothesis was rejected for the occurrence process where ($\boldsymbol{\theta}_1'$) were estimated from the records with 1800 hours starting time and ($\boldsymbol{\theta}_1$) were estimated from the record with 0700 starting time. For Rapid City the null hypothesis could not be rejected at the 5% level for either the Markov chain process or the ME model. Thus it appears that the time of day definition can account for some of the variability in the parameter fields but is probably not the sole cause. It should be noted that this method of determining the effects of observation time on the rainfall process does not account for the effects of evaporation from the rain gage.

The effects of evaporation and possibly other methodological factors can be examined by dividing the stations into two groups: those with observation time at midnight and in the morning hours, and those with observation times from 1700 to 1900. Lead was omitted from the analysis because of major elevation effects. An analysis of the data in Table 1 reveals that the average of the annual precipitation recorded at the eight stations read at midnight or morning is not significantly different from the average at the 11 stations with the afternoon observation time (students $t = 0.3275$), but the

TABLE 8. Parameter Identifiability Statistics: Pierre

Fourier Coefficients	Theoretical Mean	Observed Mean	Standard Deviation	Coefficient of Variation
		P_{00}		
γ_{10}	0.8202	0.8213	0.0036	0.0043
C_{11}	0.0681	0.0642	0.0042	0.0657
ϕ_{11}, rad	3.0101	3.0378	0.0891	0.0293
C_{12}	0.0239	0.0195	0.0035	0.1772
ϕ_{12}, rad	0.5885	0.6080	0.2843	0.4676
C_{13}	0.0124	0.0072	0.0094	1.3156
ϕ_{13}, rad	−0.5478	−0.4449	0.3636	0.8173
		P_{10}		
γ_{20}	0.6170	0.6165	0.0122	0.0197
C_{21}	0.0586	0.0709	0.0134	0.1897
ϕ_{21}, rad	−2.9438	−2.9099	0.1425	0.0490
C_{22}	0.0466	0.0388	0.0158	0.4059
ϕ_{22}, rad	1.9987	1.9869	0.3582	0.1803
C_{23}	0	*	*	
ϕ_{23}, rad		*	*	
		Alpha		
γ_{30},	0.4462	0.4348	0.0213	0.0489
		Beta		
γ_{40}, mm	0.8156	0.7782	0.0356	0.0457
C_{41}, mm	0.4648	0.4806	0.0470	0.0978
ϕ_{41}, rad	−0.7626	−0.8032	0.0594	0.0739
C_{42}, mm	0.0991	0.0345	0.0452	1.3088
ϕ_{42}, rad	−0.8102	−1.1318	0.3622	0.3200
C_{43}, mm	0	*	*	
ϕ_{43}, rad		*	*	
C_{44}, mm	0	0.0081	0.0256	3.1562
ϕ_{44}, rad		−0.9318	†	
		Delta		
γ_{50}, mm	7.7158	7.6682	0.2052	0.0268
C_{51}, mm	4.2088	0.1135	0.3592	3.1633
ϕ_{51}, rad	−0.6764	−3.6167	†	
C_{52}, mm	0.6553	*	*	
ϕ_{52}, rad	0.1056	*	*	
C_{53}, mm	0	0.0881	0.2786	3.1614
ϕ_{53}, rad		2.0320		

*Harmonic not significant at 0.05 level.
†Harmonic identified as being significant for only one simulation.

of two wet days and by the evaporation of small precipitation amounts before the gage is read.

Both of these factors could affect the number of wet days and the distribution of rainfall amounts. Differences should be most apparent on days with small rainfall amounts, so if the threshold is raised from 0.254 mm (0.01 inch) to a higher level, spatial variability of parameters should decrease. To examine this factor, we analyzed a subset of six stations: Aberdeen, Pierre, Pollock, Redfield, Gettysburg, and Onida. The MCME Fourier coefficients were identified for three thresholds, 0.254 mm, 1.27 mm, and 2.54 mm. As the threshold, T, is raised from 0.254 to 1.27 mm, the variances of the coefficients γ_{10} and γ_{20} are significantly reduced ($F > 5.05$ $df_1 = df_2 = 5$), while the coefficients γ_{30}, γ_{40}, and γ_{50} are not significantly changed (see Table 9). As the threshold is raised to 2.54 mm, the variances of the coefficients γ_{10} and γ_{20} are significantly reduced as compared to the coefficients for $T = 0.254$ mm, but the variance of coefficient γ_{50} shows a significant increase, possibly reflecting the reduced number of wet days at the higher threshold. The effect of the three thresholds on the

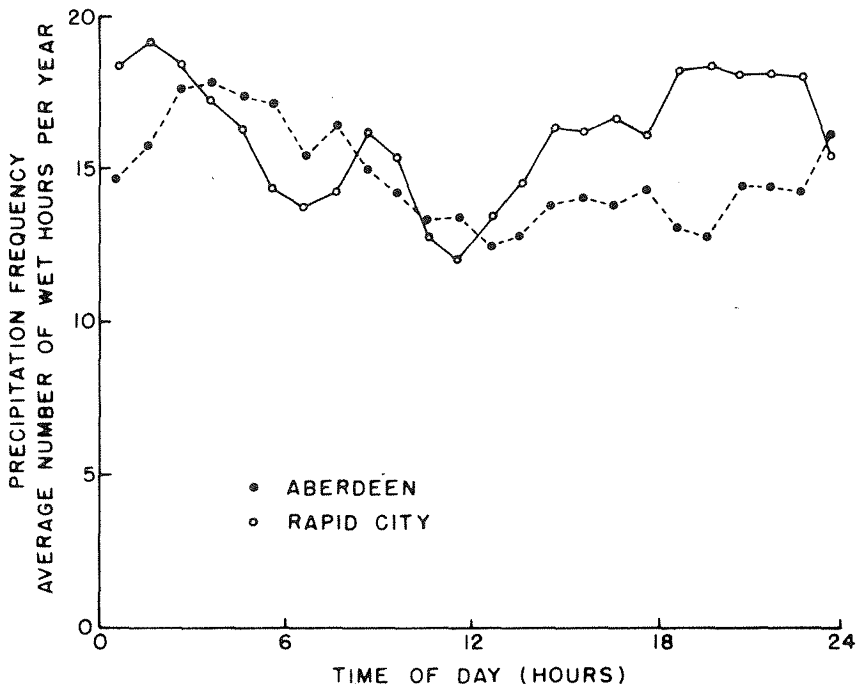

Fig. 8. Hourly frequency of precipitation occurrence.

parameter variability is introduced by methodological differences that affect the small precipitation amounts. Observation time appears to be a significant factor, and its effect can be attributed to the diurnal variability of precipitation occurrence and evaporation of small amounts of rain, so that gages serviced in the P.M. show smaller numbers of wet days than those serviced at midnight or in the morning. Although

TABLE 9. Effect of Threshold on Variability of Fourier Coefficients for Six Stations

Parameter	Coefficient	Mean	Standard Deviation	F
	Threshold = 0.254 mm (0.01 inches)			
p_{00}	γ_{10}	0.8462	0.0252	
p_{10}	γ_{20}	0.6774	0.0464	
α	γ_{30}	0.5304	0.0971	
β	γ_{40}	2.5343	1.5166	
δ	γ_{50}	9.4856	1.0688	
Number of wet days		66.284	12.064	
	Threshold = 1.27 mm (0.05 inches)			
p_{00}	γ_{10}	0.8887	0.0085	8.79*
p_{10}	γ_{20}	0.7507	0.0147	9.96*
α	γ_{30}	0.6677	0.0610	2.53
β	γ_{40}	4.4770	0.6952	4.75
δ	γ_{50}	11.1693	1.1219	1.10
Number of wet days		47.080	3.59	
	Threshold = 2.54 mm (0.10 inches)			
p_{00}	γ_{10}	0.9126	0.0067	14.17*
p_{10}	γ_{20}	0.7885	0.0118	15.46*
α	γ_{30}	0.6337	0.1482	2.33
β	γ_{40}	5.1724	0.8176	3.44
δ	γ_{50}	11.8786	2.4196	5.12*
Number of wet days		35.699	2.533	

much of this variability, particularly in the occurrence process, can be removed by using a threshold higher than 0.254 mm, it is not clear how one could estimate the parameters for the process with $T = 0.254$ mm, given the parameters for a higher threshold.

DISCUSSION

The fundamental assumption involved in mapping Fourier coefficients to provide a concise regional description of daily precipitation is that the model parameters, as represented by the coefficients, vary smoothly over the region. This is, of course, the assumption we make when we draw isolines of mean annual precipitation, so it is intuitively appealing. An examination of the parameter maps (Figures 2 through 6), the semivariograms for the mean Markov chain parameters (Figure 7), and the results of the comparisons of interpolation procedures, however, shows that there is a substantial variation of parameters in distances of the order of 50 to 100 km. An important question raised by this investigation is, "How much of this variability is real, representing true differences in the precipitation regime, and how much is spurious, introduced by measurement errors, operator bias, time of reading the gages, and modeling errors?" Our investigation revealed that much of the observed variability could be attributed to methodological differences which affect the small rainfall amounts and appear to be most consistently related to time of observation. However, parameter identifiability and sampling errors also contribute. Real mesoscale differences in the precipitation regime certainly exist, but this contribution to spatial variability of parameters is unknown.

Although we have shown that parameters estimated by four interpolation schemes were statistically different from parameters identified for four test stations, it is possible that, in some cases, information derived from precipitation sequences simulated using the interpolated parameters (for example, water

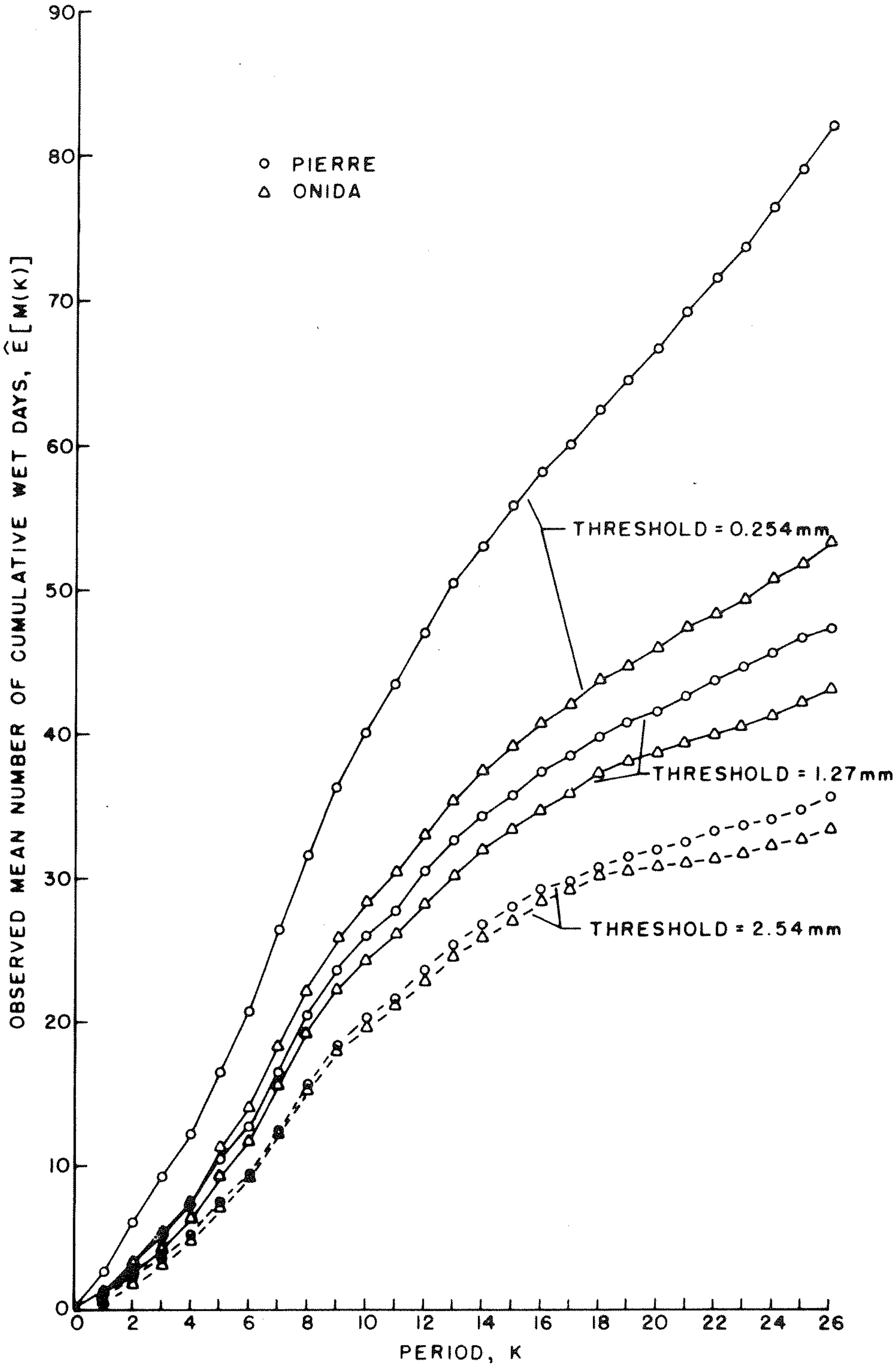

Fig. 9. Effect of threshold on expected number of wet days for adjacent stations (Pierre and Onida).

ences are caused by observation times. It should also be noted that the observation time for three of the four test stations is at 1800. The interpolated parameter estimates for these stations were based upon stations which included different observation times. Thus it appears that significant inconsistencies in data are present and may cause problems of unknown severity in regionalizing procedures and testing.

SUMMARY AND CONCLUSIONS

Fourier series are used to describe the seasonal variation of the five parameters for a stochastic model of daily precipitation utilizing the Markov chain–mixed exponential (MCME) model. Numerical maximum likelihood techniques were used to estimate the Fourier coefficients, and a likelihood ratio test of the 0.05 level was used to test the significance of each harmonic. The weighting parameter, α, in the mixed exponential distribution, was constrained to be a constant throughout the year.

A concise description of seasonal variations of parameters for the state of South Dakota has been obtained by using from 15 to 27 coefficients. This procedure provides much more information than, for instance, a listing of the monthly mean precipitation and requires only a few more parameters. Spatial variability of the mean of each parameter has been illustrated by mapping isopleths.

Semivariograms calculated for the mean Markov chain parameters γ_{10} and γ_{20} showed a "nugget" effect. The source of the large nugget variance was examined. We found that much of the observed spatial variability in parameters, at distances of 100 km or less, may be attributed to real differences in the precipitation regime and to inconsistencies in the records due to methodological differences affecting small precipitation amounts. Time of observation appears to be an important factor, but parameter identifiability and sampling error also contribute. This suggests that precipitation records proposed for use in regional parameter mapping must be carefully screened to ensure consistency of data.

The MCME parameters for four test stations were more closely estimated by arithmetic averages of the parameters of six nearby stations than by three other interpolation techniques, including nearest neighbor, spline fitting, and linear interpolation. This finding is consistent with the variograms calculated for the mean Markov chain parameters. This suggests that this estimation procedure is superior to the commonly used practice of transposing precipitation records rather long distances (other factors, such as length of record, being equal) and that the more complex interpolation procedures, such as kriging or spline fitting, are not justified. We also found that the interpolated parameters for the four test stations were significantly different from parameters identified from precipitation records.

Geographical barriers obviously affect the precipitation climatology. Therefore the application of this model to a regional description is not recommended in mountainous regions.

Acknowledgments. Contribution of the Agricultural Research Service, U.S. Department of Agriculture (USDA), in cooperation with the Colorado State University Experiment Station. J. Roldán received financial support from the Ministry of Education, Spain, for postdoctoral training and research at Colorado State University and the Agricultural Research Service, USDA, Fort Collins, Colorado. T. Econopouly assisted with the computer programming. We thank C. L. Hanson, USDA, ARS, and W. F. Lytle, South Dakota State University, for their assistance in obtaining the basic data for this study.

REFERENCES

Akaike, H., A new look at the statistical model identification, *IEEE Trans. Automat. Contr., AC-19*(6), 716–723, 1974.

Barger, G. L., R. H. Shaw, and R. F. Dale, Chances of receiving selected amounts of precipitation in north central region of the United States, report, Iowa State Univ. of Sci. and Technol., Ames, 1959.

Buishand, T. A., Stochastic modeling of daily rainfall sequences, *Rep. 77-3,* 211 pp., Mededelingen Landbouwhogeschoul, Wageningen, Netherlands, 1977.

Coe, R., and R. D. Stern, Fitting models to daily rainfall, *J. Appl. Meteorol., 12*(7), 1024–1031, 1982.

Creutin, J. D., and C. Obled, Objective analyses and mapping techniques for rainfall fields: An objective comparison, *Water Resour. Res., 18*(2), 413–431, 1982.

Duchon, J., Interpolation des fonctions de 2 variables suivant le principe de la flexion des plaques minces, *Rev. Automat. Inf. Rech. Oper., 10*(12), 5–12, 1976.

Feyerherm, A. M., and L. D. Bark, Statistical methods for persistent precipitation patterns, *J. Appl. Meteorol. 4*, 320–328, 1965.

Feyerherm, A. M., L. D. Bark, and W. C. Burrows, Probabilities of sequences of wet and dry days in South Dakota, *North Cent. Reg. Res. Publ., 161*, Kans. State Univ., Manhattan, 1965.

Fletcher, R., Fortran subroutines for minimization by quasi-Newton methods, *Rep. AERE-R7125*, Atomic Energy Research Establishment, Harwell, England, U. K., 1972.

Haan, C. T., D. M. Allen, and J. O. Street, A Markov chain model for daily rainfall, *Water Resour. Res., 12*, 443–449, 1976.

Hoel, P. G., *Introduction to Mathematical Statistics*, John Wiley, New York, 1971.

Hughes, J. P., and D. P. Lettenmaier, Data requirements for kriging: Estimation and network design, *Water Resour. Res., 17*(6), 1641–1650, 1981.

Jones, J. W., R. F. Colwick, and E. D. Threadgill, A simulated environmental model of temperature, evaporation, rainfall, and soil moisture, *Trans. ASAE, 15*(2), 366–372, 1972.

Katz, R. W., Precipitation as a chain dependent process, *J. Appl. Meteorol., 16*, 671–676, 1977.

Mielke, P. W., and E. S. Johnson, Three-parameter kappa distribution maximum likelihood estimates and likelihood ratio tests, *Mon. Weather Rev., 101*(9), 701–707, 1973.

Paihua, L., and F. Utreras, Un ensemble de programes pour l'interpolation des fonctions par des splines de type plaque mince, *Res. Rep. 140*, Lab. de Math. Appl., Grenoble, France, 1978.

Richardson, C. W., and D. A. Wright, WGEN: A model for generating daily weather variables, *U.S. Agric. Res. Serv., ARS-8*, 1–83, 1984.

Roldán, J., and D. A. Woolhiser, Stochastic daily precipitation models, I, A comparison of occurrence processes, *Water Resour. Res., 18*(5), 1451–1459, 1982.

Skrivan, J. A., and M. R. Karlinger, Semi-variogram estimation and universal kriging program, U.S. Geol. Surv., Water Resour. Div., Tacoma, Wash., 97 pp., 1979.

Smith, R. E., and H. A. Schreiber, Point process of seasonal thunderstorm rainfall, 1, Distribution of rainfall events, *Water Resour. Res., 9*(4), 871–884, 1973.

Smith, R. E., and H. A. Schreiber, Point process of seasonal thunderstorm rainfall, 2, Rainfall depth probabilities, *Water Resour. Res., 10*(3), 418–423, 1974.

Stern, R. D., Analysis of daily rainfall at Samaru, Nigeria, using a simple two-part model, *Arch. Meteorol. Geophys. Bioklimatol. Ser. B, 28*, 123–135, 1980*a*.

Stern, R. D., The calculation of probability distributions for models of daily precipitation, *Arch Meteorol. Geophys. Bioklimatol. Ser. B, 28*, 137–147, 1980*b*.

Todorovic, P., and D. A. Woolhiser, Stochastic model of *n*-day precipitation, *J. Appl. Meteorol., 14*(1), 17–24, 1975.

Woolhiser, D. A., and G. G. S. Pegram, Maximum likelihood estimation of Fourier coefficients to describe seasonal variations of parameters in stochastic daily precipitation models, *J. Appl. Meteorol., 18*, 34–42, 1979.

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

Woolhiser, D. A., E. W. Rovey, and P. Todorovic, Temporal and spatial variation of parameters for the distribution of *n*-day precipitation, in *Floods and Droughts, Proceedings of the Second International Symposium in Hydrology*, edited by E. F. Schulz, V. A. Koelzer, and K. Mahmood, pp. 605–614, Water Resources Publications, Fort Collins, Colo., 1973.

J. Roldán, Catedra de Hidraulica General y Agricola E. T. S. I. Agronomos, Universidad de Cordoba, Apartado 3.048, Cordoba, Spain.

D. A. Woolhiser, U.S. Department of Agriculture, Agricultural Research Service, 2000 East Allen Road, Tucson, AZ 85719.

(Received August 6, 1985;
revised January 7, 1986;
accepted January 28, 1986.)