FLOODS AND DROUGHTS

Reprinted from the
Proceedings of the
SECOND INTERNATIONAL SYMPOSIUM
IN HYDROLOGY
September 11-13, 1972

TEMPORAL AND SPATIAL VARIATION OF PARAMETERS FOR THE DISTRIBUTION OF N-DAY PRECIPITATION

By

D. A. Woolhiser
Research Hydraulic Engineer, ARS-USDA

E. Rovey
Graduate Research Assistant
Colorado State University

P. Todorovic
Associate Professor of Civil Engineering
Colorado State University

Fort Collins, Colorado

Abstract

A proposed stochastic model of n-day rainfall is structured such that the number of rainy days in an n-day period, N_n , is described by a Markov Chain and the amount of rainfall per day $\geq$ 0.01 inch is exponentially distributed. Parameters for this model were estimated for 7- and 21-day intervals for 11 eastern Colorado stations. The time variations of the parameters were fitted using a finite Fourier series and the spatial variation was accounted for by maps of the annual mean value. Such mean value maps along with a set of normalized Fourier coefficients provide a regionalized model of n-day precipitation. This model has several important advantages over the currently used gamma distribution and provides equally good fits to observed distributions.

Introduction

The distribution function of the total amount of rainfall in n days has many practical applications in engineering and agriculture. The approach commonly used is to transform the data to obtain approximate normality or to fit a gamma distribution to the observations (see Skees and Shenton [1] for a recent review of these techniques). Although adequate fits can be obtained, the parameters of the distributions vary with the number of days in an unpredictable manner.

Todorovic and Woolhiser [2] proposed the following general form of the distribution function for n-day rainfall:

$$F_n(x) = P\{N_n = 0\} + \sum_{\nu=1}^{n} P\{X_\nu \leq x\} P\{N_n = \nu\} \quad , \qquad (1)$$

where N_n is the number of rainy days in an n-day period and X_ν is the sum of ν daily rainfall amounts, i.e.,

$$X_\nu = \sum_{i=1}^{\nu} \xi_i \quad . \tag{2}$$

In this formulation ξ_i's are independent, identically distributed random variables with finite mean and variance.

One of the special cases considered by Todorovic and Woolhiser was the Markov Chain-Exponential Model where the distribution function $P\{X_\nu \leq x\}$ was computed for an exponential distribution function of daily rainfall:

$$H(x) = 1 - e^{-\lambda x} \quad . \tag{3}$$

The distribution function $P\{N_n = \nu\}$ was developed based on the assumption that the sequence of random variables η_1, $\eta_2, \ldots \eta_n$ is a Markov Chain such that for all $\nu = 1, 2 \ldots n$

$$P\{\eta_\nu = 1 | \eta_{\nu-1} = 0\} = q_o \quad ,$$

$$P\{\eta_\nu = 1 | \eta_{\nu-1} = 1\} = q_1 \quad ,$$

where

$$\eta_j = \begin{cases} 1 \text{ if } j^{th} \text{ day is wet} \\ 0 \text{ if } j^{th} \text{ day is dry} \end{cases} \quad .$$

This is the process for number of rainy days developed by Gabriel [3] and applied by Gabriel and Neuman [4].

The distribution function of n-day rainfall for the Markov Chain-Exponential model can be written as follows (Todorovic and Woolhiser, 1971):[2]

$$F_n(x) = \{(1 - q_o) - R\, d\}\, (1 - q_o)^{n-1}$$

$$+ \sum_{\nu=1}^{n} \{R\, \phi_{\cancel{o}\,1}(\nu,n) + (1-R)\, \phi_{\cancel{1}\,o}(\nu,n)\} \frac{\lambda^\nu}{\Gamma(\nu)} \int_o^x u^{\nu-1} e^{-\lambda u} du \ , \tag{5}$$

where $R = P\{\eta_o = 1\}$, η_o is the random variable which refers to the day preceding the first day of the n-day period, $d = q_1 - q_o$, $\phi_o(\nu,n) = P\{N_n = \nu | \eta_o = 0\}$, $\phi_1(\nu,n) = P\{N_n = \nu | \eta_o = 1\}$ and λ is the parameter of the exponential distribution.

The model specified by Eq. 5 has five parameters which can be readily interpreted in terms of the length of period, the nature of the counting process for the number of rainy days or the distribution of daily rainfall amounts. In this paper we investigate the spatial and temporal variation of these parameters for eleven rainfall stations in eastern Colorado for periods of 7 and 21 days.

Computation of Parameter Values

Daily precipitation data for the eleven eastern Colorado stations shown in Table 1 were available on magnetic tape, and had been previously utilized in a regional publication (Heermann, Finkner and Hiler [5]).

Table 1. Precipitation Data Utilized

Station	Elevation (ft)	Mean Annual Precipitation (in.)	Beginning Year of Record	Length of Record (yrs)
Akron	4582	16.75	1930	36
Burlington	4171	16.35	1931	37
Denver	5221	12.89	1931	37
Fort Collins	5004	14.19	1931	37
Julesburg	3469	16.32	1931	36
Lamar	3635	14.20	1931	37
Limon	5560	14.94	1931	37
Pueblo	4639	11.84	1931	37
Rocky Ford	4178	12.31	1931	37
Trinidad	6030	14.90	1925	34
Two Buttes	4075	13.85	1929	36

Parameter values were computed for weekly (7-day) periods beginning March 1; February 28 and 29 were included in the 52nd week. The Markov Chain parameters $\hat{q}_0$ and $\hat{q}_1$ were weekly averages obtained from a western regional publication (Heermann, Finkner and Hiler [5]). The parameter $\hat{R}$ was taken as the number of wet days in the previous 7-day period divided by the number of days. The parameter $\hat{\lambda}$ was taken as the inverse of the sample mean of all daily precipitation amounts ≥ 0.01 inch. These parameters were also estimated for a 21-day period. The temporal variation of weekly $\hat{R}$, $\hat{q}_0$, $\hat{q}_1$ and $1/\hat{\lambda}$ is shown for Fort Collins in Fig. 1. The 21-day mean daily precipitation for the mountain front stations, Fort Collins, Denver, Pueblo and Trinidad, typically showed two maxima: one in May and a broad plateau beginning in July and continuing through September. The plains stations showed one maximum in late July or early August.

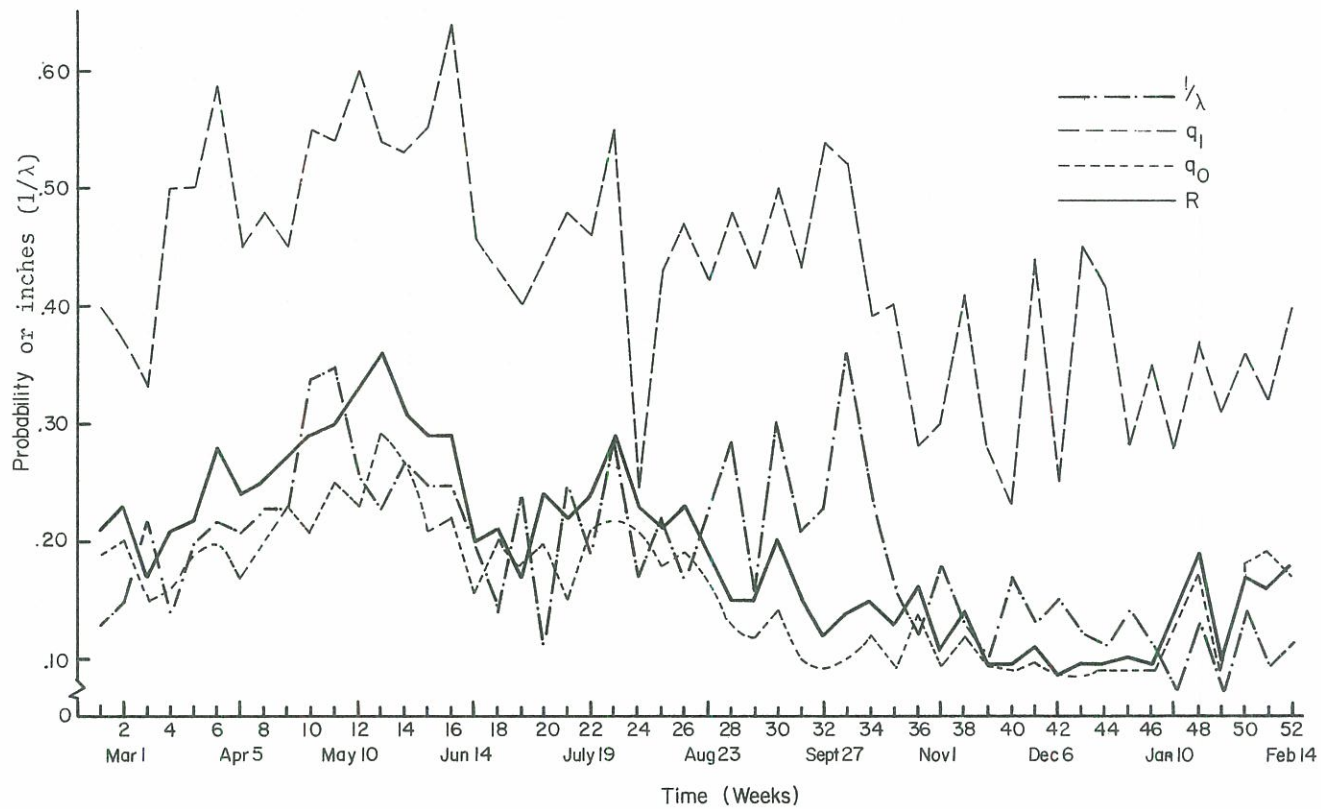

Figure 1. Temporal variations for weekly sample values of parameters for Fort Collins, Colorado.

The 7-day parameter values for each station were fitted with a finite Fourier series:

$$f(x) = a_o + \sum_{n=1}^{12} \left(a_n \cos \frac{n\pi x}{26} + b_n \sin \frac{n\pi x}{26}\right) , \qquad (6)$$

$$x = 1, 2, \ldots 52$$

All coefficients were then normalized by dividing by a_o and the stations were divided into two groups exhibiting similar time variations in the parameters: (1) the stations near the foothills, and (2) the plains stations. The normalized coefficients for each group were averaged; if the average coefficient was < 0.05, it was dropped. The normalized coefficients > 0.05 are shown in Table 2, and the mean values a_o are plotted on

Table 2. Averaged, Normalized Fourier Coefficients

		Mountain Front		Plains Stations	
		7-day	21-day	7-day	21-day
λ	a_1	.280	.312	.371	.418
	a_2	.090	.174	*	.095
	a_4	*	*	*	.056
	b_1	-.248	-.191	-.319	-.227
	b_2	-.228	-.173	-.140	-.122
R	a_1	-.087	-.170	.177	-.277
	a_2	*	*	-.052	*
	a_4	.072	.083	.068	.054
	b_1	.409	.386	.526	.486
	b_2	*	*	-.063	*
	b_5	.073	.059	.061	.055
q_o	a_1	-.106	-.191	-.159	-.252
	a_2	.070	*	*	*
	a_4	.094	.110	.063	*
	b_1	.402	.374	.508	.470
	b_2	.081	*	-.093	-.080
	b_5	.068	*	.071	.052
q_1	a_1	*	-.052	-.100	-.132
	a_2	*	*	*	-.065
	a_4	*	*	*	.101
	a_6	*	*	-.053	*
	b_1	.163	.153	.254	.231
	b_2	*	*	.091	.084
	b_3	*	*	-.065	*

* < 0.05

the maps of Fig. 2. The maps of mean parameter values in Fig. 2 and the normalized Fourier coefficients in Table 2 provide a regionalized estimate of parameters for the Markov Chain-Exponential Model for 7-day periods. Similar procedures were used to obtain regionalized parameters for 21-day periods.

Comparison of Computed and Observed Distributions

The theoretical distributions of 7- and 21-day rainfall amounts were computed using the computer program presented by Todorovic and Woolhiser [2] with parameters estimated from the data and also with the regionalized parameters. Three periods were selected for each station: spring, summer and fall. Figure 3 shows the best and the worst fit out of a sample of 15 distributions for 7-day and 21-day rainfall amounts based upon the maximum deviation between observed and computed distribution functions with regionalized parameters:

$$D = \max \left| \hat{F}_n(x) - F_n(x) \right|$$

where $\hat{F}_n(x)$ is the empirical distribution and $F_n(x)$ is the theoretical distribution with regionalized parameters. Precipitation probabilities obtained by fitting the incomplete gamma distribution to the observational data and smoothing by means of a 3-week moving average in which double weight was given to the week under consideration (Gifford, Ashcroft and Magnuson [6]) are also plotted in Fig. 3. The maximum deviations between the observed distribution and the incomplete gamma were also measured and compared with the Markov Chain-Exponential models. Based on this criterion, the Markov Chain-Exponential model with regionalized parameters gave the best fit for seven distributions, while the gamma gave the best fit for six cases. The Markov Chain-Exponential model with observed parameters gave the best fit for four distributions (there are two ties). It is difficult to make comparisons because of the different techniques used. The regionalized parameter model involves smoothing in both space and time, the gamma utilizes time smoothing by means of a moving average, and the Markov Chain-Exponential with parameters estimated for that period involves only averaging within the period. Caskey [7] demonstrated that the Markov Chain model was quite satisfactory for rainfall occurrence at Denver. We used the chi-square test for goodness of fit of the exponential distribution to daily rainfall amounts. The exponential distribution is quite satisfactory for daily rainfall sampled within a 7-day interval, but the null hypothesis could be rejected at the 0.95 level for several samples obtained within a 21-day interval. Because of our method of estimating parameters for the Markov Chain-Exponential model, a chi-square test is not appropriate. However, on the basis of the maximum deviation and graphical comparisons with the gamma distribution, it appears that the Markov Chain-Exponential

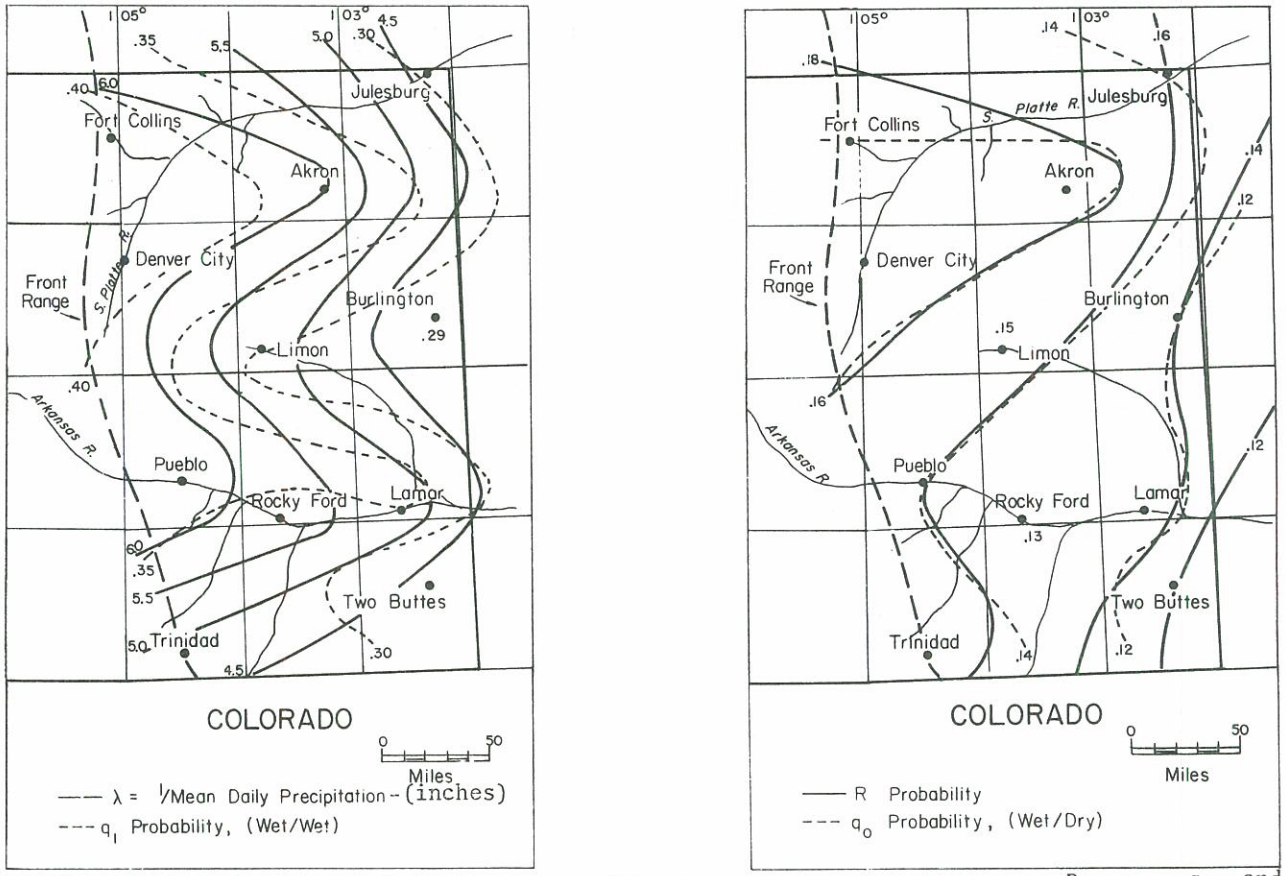

Figure 2. Spatial variations of the Fourier coefficient a_o for the parameters R, q_o, q_1 and λ.

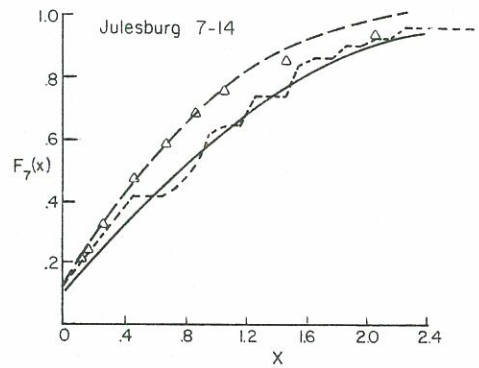

Figure 3a. Worst fit for 7-day rainfall.

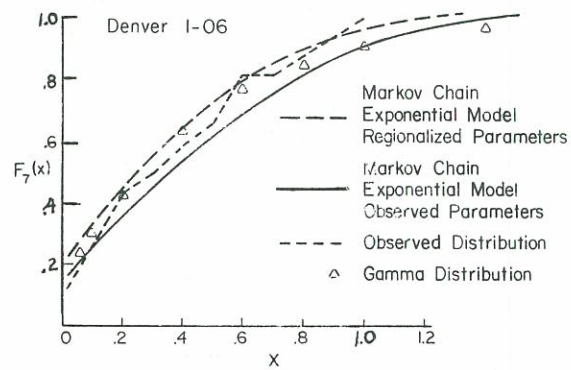

Figure 3b. Best fit for 7-day rainfall.

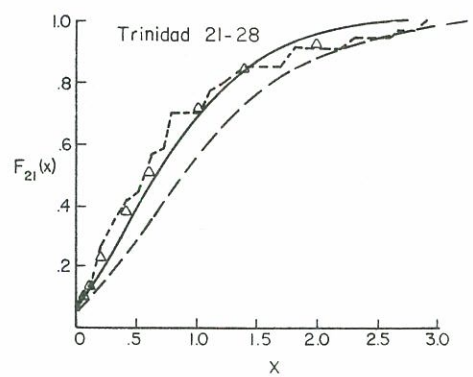

Figure 3c. Worst fit for 21-day rainfall.

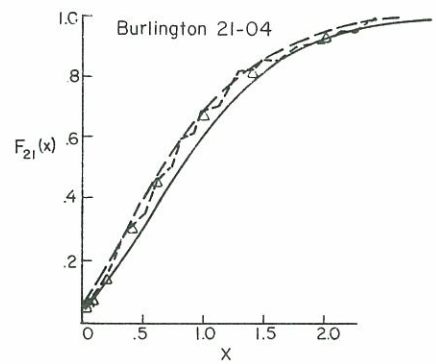

Figure 3d. Best fit for 21-day rainfall.

model with the regionalized parameters fits the observed data at least as well as the gamma distribution. From Fig. 2 it is obvious that the annual mean parameters $\bar{q}_o$ and $\bar{R}$ are highly correlated, with:

$$\bar{q}_o \approx \bar{R} - 0.02 \quad .$$

Furthermore, there seems to be some correlation between $\bar{q}_1$ and $\bar{\lambda}$ which should be further investigated.

Discussions and Conclusions

We have presented a convenient method for estimating the spatial and temporal variation of parameters for the Markov Chain-Exponential Rainfall model proposed by Todorovic and Woolhiser. The Markov Chain-Exponential model has the following advantages over the currently used gamma distribution: (1) Parameters of the model are easily interpreted in terms of the climatology of the area; (2) parameters appear to be regionally smooth in eastern Colorado, suggesting that the model can be regionalized for other areas with similar topographic variations; and (3) the model enables the computation of the distribution function of n-day rainfall for any value of n, subject only to the condition that the other four parameters do not vary significantly within the period.

The Markov Chain process has been shown to be quite successful in describing the number of rainy days in an n-day period (Gabriel and Neuman [4], Caskey [7], Feyerherm and Bark [8]). Some improvement in the general model specified by Eq. 1 might be obtained by considering a distribution other than the exponential for daily rainfall amounts ξ_1, $\xi_2 \ldots \xi_k$. Any improvement, however, would probably involve at least one additional parameter, making the regionalizing method more difficult.

References

[1] Skees, P., and L. R. Shenton, Comments on the Statistical Distribution of Rainfall Per Period Under Various Transformations, Proc. of the USDA-IASPS Symposium on Statistical Hydrology, Tucson, Arizona, (1971).

[2] Todorovic, P., and D. A. Woolhiser, Stochastic Model of Daily Rainfall, Proc. of the USDA-IASPS Symposium on Statistical Hydrology, Tucson, Arizona, (1971).

[3] Gabriel, K. R., The Distribution of the Number of Successes in a Sequence of Dependent Trials, Biometrika, V. 96, pp. 959-960, (1959).

[4] Gabriel, K. R., and J. Neuman, A Markov Chain Model for Daily Rainfall in Tel Aviv, Quar. Jour. of Royal Met. Soc., Vol. 88, pp. 90-95, (1962).

[5] Heermann, Dale F., Morris D. Finkner and Edward A. Hiler, Probability of Sequences of Wet and Dry Days for Eleven Western States and Texas, Excerpt from Tech. Bulletin 117, Colorado State University Experiment Station, Fort Collins, (1971).

[6] Gifford, R. O., G. L. Ashcroft and M. D. Magnuson, Probability of Selected Precipitation Amounts in the Western Region of the United States, Section for Colorado, T-8 Agr. Exp. Sta. Univ. of Nevada, (October 1967).

[7] Caskey, James E., Jr., A Markov Chain Model for the Probability of Precipitation Occurrence in Intervals of Various Length, Monthly Weather Review, pp. 298-301, (June 1963).

[8] Feyerherm, A. M., and L. Dean Bark, Statistical Methods for Persistent Precipitation Patterns, Jour. Appl. Meteorol., Vol. 4, pp. 320-328, (June 1965).

Purchased by
USDA Agric. Research Service
for Official Use.