Journal of Agrometeorology 4 (1) : 65 - 73 (June 2002)

Stochastic model for weekly rainfall of Junagadh

R. SUBBAIAH AND D. D. SAHU[1]
College of Agricultural Engineering and Technology, [1]Department of Agronomy,
Gujarat Agricultural University, Junagadh - 362001

ABSTRACT

A model for the weekly monsoon season rainfall is presented based on the assumption that weekly rainfall in the season is a first order Markovian process. Comparison between the historical and synthetic series shows that the two are statistically comparable with respect to measures of central tendency, dispersion and distribution.

Key Words : Rainfall model, Gamma distribution, Markovian chain

Irrigation scheduling and design and operation of the irrigation systems are directly dependent on the input of rainfall. The three main characteristics of rainfall i.e. the amount, the intensity, and frequency have wide spatial and temporal variations in arid and semiarid regions. Therefore, accurate prediction of the rainfall over the entire season and at the peak period, as well as its distribution over the growing season, are of great importance for the rational design and management of irrigation systems.

Planning and operation of canal irrigation schemes is generally based upon either the weekly, fortnightly or monthly values of rainfall estimated at various recurrence intervals from historical series. However probabilistic models of rainfall are useful in planning irrigation projects rather than in their operation. Attaining high water use efficiency and agricultural production remain elusive goals in the canal command areas. If suitable models for forecasting rainfall are available, the efficiency of operation of large irrigation schemes can substantially be enhanced.

Numerous models are available to generate the daily rainfall using Markov chains (Stern and Coe, 1984), and alternating renewal process (Buishand, 1977; 1978; Roldan and Woolhiser, 1982; and Tskarais *et al.,* 1984). Waymire and Gupta (1981) presented a detailed review of the above models. Both type of models need large historical data sets and require estimation of a large number of parameters. But the real time application of the above models is limited because the state of the system needs to be updated regularly and the complex two stage models make this task difficult.

A closer look at the literature along these lines reveals that although many mathematical models have been proposed on the structure of rainfall, there is no unified mathematical approach to modeling rainfall. In part, this difficulty stems from the considerable space-time variability of rainfall and non-availability of appropriate mathematical tools designed to exploit the clustering dependence, which the rainfall phenomenon seems to exhibit.

The week is a practical time step for scheduling irrigation's on a rotational basis in large canal schemes and planning inter-culturing operations in crops. Hence this paper aims at developing a simple weekly rainfall model useful for real time irrigation scheduling.

MATERIALS AND METHODS

During the monsoon season the probability that a wet week is followed by a wet week is much higher than by a dry week. The weekly rainfall process is therefore considered to be represented by a first order Markovian process. This process has the property that the state of the system in any week depends on the state of the system in the previous week only and not on the earlier history of the system. The probability that a random variable X will change from state i to state j in successive time intervals is called the transition probability P_{ij}. For a first order Markov process with n discrete states,

$$P_{ij} = P\,(W_{t+1=j} \,/\, W_{t=i}) \qquad (1)$$
$$i = 1,2,3,....n;\; j = 1,2,3,....n$$

The P_{ij} from a square matrix P and are the one step transition probabilities. The probabilities P_{ij} are cumulated along each row i to yield the transition probability matrix P. The derivation of P matrix is carried out in a number of stages. First of all the number of states n is defined and the range covered by each state is fixed. Then the weekly rainfall data series are examined to find the frequency of transitions from state i to state j. These counts are then scaled so that each row adds up to unity thereby forming a stochastic matrix. The weekly rainfall series were generated using the P matrix and a random number generator.

a) The initial state of the system is presumed to be known.

b) A random number rt with same probability density function as the weekly rainfall series is obtained. For the row corresponding to state i, bilinear interpolation, the column of the transition probability matrix, which corresponds to it is determined. This is the next state to be visited, state j at time (t+1). The mid point of this state is assumed to be the rainfall for the week (t+1).

c) State j then becomes state i and i is increased by 1 for simulating rainfall of the next week.

d) We return to step (b) and repeat till the required sequence is produced.

An important requirement for the simulation model is the distribution of random numbers used to generate the synthetic data. Nair and Sudarsan (1997) have fitted normal, lognormal, log Pearson, gamma and exponential distributions to weekly rainfall data. The standard error of the distribution is selected as a criterion to identify the best-suited frequency distribution for the weekly rainfall data. The standard error of the distribution is defined as:

$$S_e = \sum_{i=1}^{n} [(x_i - y_i)^2 (n - m_j)^{-1}] \qquad (2)$$

where x_i and y_i are respectively the recorded events and the event magnitudes computed from the j^{th} probability distribution at probabilities computed from the stored ranks of x_i and m_j is the number of parameters estimated from the j^{th} distribution. They have suggested exponential distributions to 38 to 40^{th} standard weeks and gamma distributions to remaining weeks in the

Table 1 : Relative standard error for various frequency distributions of weekly rainfall.

Standard Week	Relative standard error					Best suited distribution (Col No.)
	Normal	Lognormal	Log Pearson	Exponential	Gamma	
1	2	3	4	5	6	7
26	11.2582	8.3214	5.4289	7.6248	2.3238	6
27	10.6964	8.8897	6.2874	9.3154	1.0177	6
28	10.1123	6.3258	3.3399	5.1128	1.8524	6
29	11.3258	6.2541	5.8524	8.1247	2.3648	6
30	13.5879	7.3698	5.6178	10.1123	3.3542	6
31	9.8564	5.2811	2.9874	6.1594	1.8974	6
32	9.3587	4.9987	3.0214	5.1112	1.1117	6
33	10.3214	7.6325	5.8975	4.3571	2.0214	6
34	12.3698	7.1123	4.3287	8.2697	0.1569	6
35	11.8748	3.2587	1.9874	7.0123	0.2478	6
36	13.5821	4.1128	1.5632	5.3214	0.8974	6
37	11.0321	5.1187	2.0135	9.9085	0.9635	6
38	11.0365	6.9631	8.3128	2.3487	7.8964	5
39	10.2856	7.2483	9.5412	1.8975	8.5469	5
40	13.5847	8.9856	7.9984	4.3872	9.2654	5

monsoon season (Table 1). Table 1 shows the standard errors divided by respected standard deviations for various frequency distributions. Therefore, exponential and gamma distributed random numbers were used to generate the weekly rainfall.

Exponential distribution

If the probability that an event will occur during a small time interval is very small and if the occurrence of this event is independent of the occurrence of other events, then the time interval between the occurrence of events is exponentially distributed. The probability density function of the exponential distribution is given by:

$$f(x) = \lambda \exp(-\lambda x) \quad (3)$$

In which λ is the inverse of mean weekly rainfall.

Gamma distribution

The gamma density function is given by

$$f(x) = \lambda^n x^{(n-1)} \exp(-\lambda x) / (n-1)! \quad (4)$$

Where λ and n are scale and shape parameters. Haan (1977) gave the moment estimators and maximum likelihood estimators for λ and n. Following the recommendation of Thom (1958), Greenwood and Durand

(1960) maximum likelihood estimator's approach is adopted. Bowman and Sherton (1970) alleviated the bias resulting from the Greenwood and Durand (1960) maximum likelihood approach.

Generation of gamma variates

The rainfall data was generated for a period that equaled five times the length of record for 26th to 40th weeks. The rainfall state of the 25th week (t=1) was assumed to be the same as that of corresponding historical rainfall series. The data were generated using gamma distributed random numbers. Several gamma random variate-generating techniques are available (Cheng, 1977; Phein and Ruksaslip, 1981; Srikanthan and MacMohan, 1983). From these techniques the Wilson-Hilferty transformation (Srikanthan and MacMohan, 1983) was adopted. According to Wilson-Hilferty transformation the approximately gamma distributed random variable is expressed as

$$\epsilon_i = 2\,[\{1+(g_k\, n_i /6) - (g^2k /36)\}^3 - 1] / g_k \quad (5)$$

$$g_k = (1 - r^3\, v_k / (1 - r^2)^{1.5} \quad (6)$$

where, ϵ_i = gamma random number with mean zero, variance one and skewness g_k; r = lag one correlation coefficient; V_k = skewness coefficient for weekly rainfall; and n_i = standardized, normal, random, variate.

The exponentially distributed random numbers were generated using Shook and Highland (1969) approach. To simulate an exponential distribution call the mean (θ) and use the inverse transformation of the density function:

$$X = -\theta \, \ln(RN) \quad (7)$$

where, X = generated exponential number; $\theta = 1/\lambda$ and RN = randomly selected number in the interval 0 to 1.

RESULTS AND DISCUSSION

Average weekly rainfall data of 40 years (1958-97) recorded by rain gauge station at Gujarat Agricultural University, Junagadh, are utilized as the input data to the model for validation. The rainfall series was tested for homogeneity using the Kruskal Wallis test and Mann Whitney test. In each case the sample was found to be homogeneous. The means, standard deviations, coefficient of variation, coefficient of skewness and kurtosis coefficient of the weekly rainfall in the monsoon season are presented in Table 2. More than 95 % of the annual rainfall in the study area is confined to monsoon season only. The onset and cessation weeks of monsoon are 25th and 40th standard weeks respectively (Nair and Sudarsan 1997). The model development is therefore confined to 26th to 40th standard weeks of the years. The weekly rainfall series have a positive skewness coefficient (Table 2). The value λ (eq.3) in exponential distribution for 38 to 40th standard weeks are respectively 0.04, 0.083 and 0.094. The unbiased estimates of scale and shape parameters for various weeks are presented in Table 2. The goodness of fit of each probability density function was tested by the Chi-square test. The difference between the observed and estimated frequencies of rainfall at 5 percent level of significance was not significant.

Transitional probability matrix

The weekly rainfall values were grouped

Table 2 : Statistical parameters, scale and shape parameters of weekly rainfall at Junagadh.

Standard weeks	Historical series				Gamma parameters		2 Static (5% level)	KS Static (5% level)
	Mean (mm)	Cv	Cs	Ck	λ	n		
26	70.68	3.00	5.59	35.30	0.0118	0.8335	6.90	0.72
27	77.50	1.63	1.42	4.98	0.0154	1.1936	9.10	0.63
28	87.70	1.18	1.12	5.61	0.0125	1.0984	8.70	0.89
29	103.03	1.26	1.26	4.07	0.0122	1.2562	9.30	0.78
30	69.55	1.13	1.54	9.61	0.0134	0.9310	10.10	0.85
31	57.82	1.47	2.36	10.10	0.0180	1.0463	9.20	0.45
32	60.22	1.32	2.47	8.81	0.0141	0.8545	8.60	0.58
33	82.73	1.60	5.41	33.09	0.063	0.5223	7.40	0.42
34	23.16	2.85	2.80	11.62	0.0395	0.9147	10.60	0.63
35	51.73	1.49	3.90	19.21	0.0130	0.6760	6.90	0.52
36	30.47	2.06	1.95	5.72	0.0342	1.0442	11.50	0.66
37	24.67	1.32	3.14	14.74	0.0317	0.7824	7.90	0.39
38	25.63	1.74	3.10	12.86			9.50	0.49
39	12.01	1.95	1.40	3.82			8.10	0.62
40	10.60	1.46	3.20	12.83			6.30	0.85

Table 3 : Definition of rainfall states.

Rainfall State	Rainfall interval (mm)	Mid point (mm)
1	0	0
2	0-10	5
3	10.1-20	15
4	20.1-30	25
5	30.1-40	35
6	40.1-50	45
7	50.1-60	55
8	60.1-70	65
9	70.1-80	75
10	>80	100

by dividing the data into 10 equal sections. The definitions of the various states are given in Table 3. The rainfall of 26th to 40th weeks for 40 years is represented by the corresponding states depending on the rainfall interval. From this table of rainfall states, the frequencies of transitions between the states were computed (Table 4). The probability transition matrix (Table 5) was determined from this table by dividing each number of this table by sum of each row and cumulative successive numbers in each row.

Model validation

The state of the system needs to be updated with current information in real time

Table 4 : Frequencies of transitions between rainfall states.

Rainfall states	1	2	3	4	5	6	7	8	9	10
1	7	19	8	0	7	2	0	8	2	0
2	10	37	0	9	8	9	17	10	0	17
3	8	0	6	5	5	0	8	9	6	0
4	6	12	0	0	4	8	9	0	3	[illegible]
5	5	0	14	3	9	6		7	4	0
6	3	8	13	8	6	10	0	8	0	8
7	0	11	0	4	3	0	7	0	5	9
8	8	9	7	13	1	7	5	12	0	10
9	0	0	9	5	3	6	0	0	1	2
10	0	5	8	14	9	0	5	0	7	0

Table 5 : Transition probability matrix of weekly rainfall at Junagadh.

Rainfall states	1	2	3	4	5	6	7	8	9	10
1	0.13	0.49	0.64	0.64	0.77	0.81	0.81	0.96	1.00	1.00
2	0.08	0.40	0.40	0.47	0.54	0.62	0.16	0.85	0.85	1.00
3	0.17	0.17	0.29	0.40	0.51	0.51	0.68	0.87	1.00	1.00
4	0.12	0.36	0.36	0.36	0.44	0.60	0.78	0.78	0.84	1.00
5	0.08	0.08	0.33	0.38	0.54	0.65	0.80	0.92	1.00	1.00
6	0.04	0.17	0.37	0.50	0.59	0.75	0.75	0.87	0.87	1.00
7	0.00	0.28	0.28	0.38	0.46	0.46	0.64	0.64	0.76	1.00
8	0.11	0.23	0.33	0.51	0.52	0.62	0.69	0.86	0.86	1.00
9	0.00	0.00	0.34	0.53	0.65	0.88	0.88	0.88	0.92	1.00
10	0.00	0.10	0.27	0.56	0.75	0.75	0.85	0.85	1.00	1.00

operation of irrigation system, which can be attained through development of simple and easy models. Model validation is presented here to assess the model performance with respect to historical rainfall series in measures of central tendency (mean), dispersion (coefficient of variation) and distribution (Skewness). The generated mean weekly rainfall ranged from 10.6 to 103mm compared to 8.0 to 94.0 mm for that of observed series for 26[th] to 40[th] weeks. For individual weeks, the mean weekly rainfall

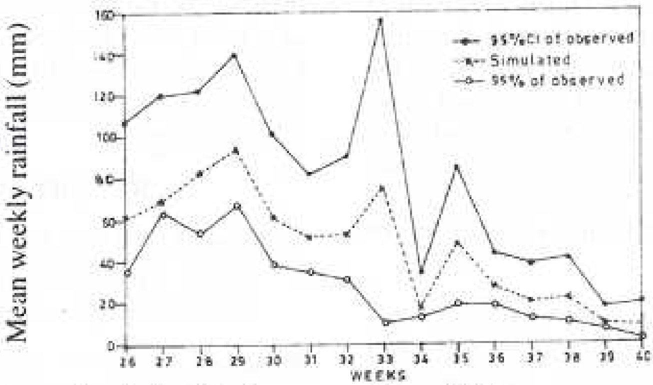

Fig 1 : Observed and simulated mean weekly rainfall.

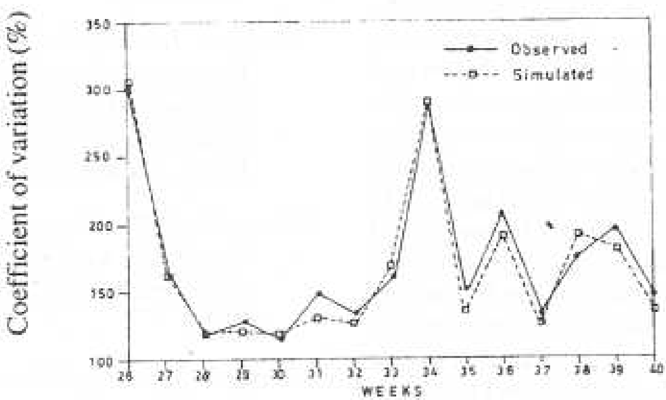

Fig 2 : Observed and simulated coefficient of variation of weekly rainfall

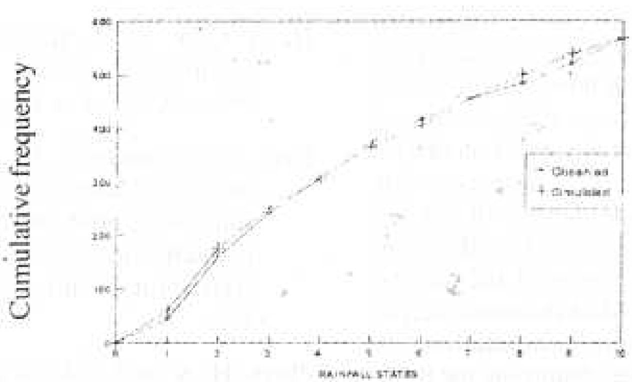

Fig 3 : Observed and simulated cumulative frequency distribution of weekly rainfall was within the 95 % confidence interval of the corresponding mean historical rainfall (Fig.1). A lower amount of rainfall of the simulated series compared to historical series was observed. This is probably due to arbitrary assignment of mid point of the rainfall state. Prediction can be improved further by considering the randomness of amount of rainfall.

The corresponding values of the coefficients of variation also were not significantly different (Fig.2). The cumulative frequency distributions of the observed and simulated series of weekly rainfall for the monsoon season as a whole were nearly identical (Fig 3). The coefficient of skewness was greater than 1.0 for all the weeks of the season for both synthetic and historical series. Comparison with observed and generated frequency distribution for individual weeks were also carried out using Chi-square and the Kolmogorov-Smirnov two sample tests to prove that the distributions of weekly synthetic and historic series are not significantly different for each week (Table 2). The difference between the weekly synthetic and historical series of rainfall at 5 percent level of significance was not significant.

Additional features can be introduced into the model to improve the simulations by increasing the states. It is however preferable to define the states according to equal probability. In this case the state boundaries are adjusted so that each state contains approximately the same number of entries. Deriving the transition matrix in this way removes a source of potential bias in the generation procedure. Modifications such as above imply only more computations of repetitive nature without disturbing the basic model framework in any way. They can therefore be included if deemed necessary at any location. The weekly rainfall model presented in this effort is therefore of a sufficiently general nature to be adopted to other locations.

REFERENCES

Buishand, T. A. 1977. Stochastic modeling of daily rainfall sequences Melded. Landbouwhogesch, Wagenengin, p.120.

Buishand, T. A. 1978. Some remarks on the use of daily rainfall models. *J. Hydrology*, 36: 295-301.

Bowman, K. O. and Sherton, L. R. 1970. Properties of estimators for the Gamma distribution report CTC-1. Union Carbde Corporation, Nuclear Division, Oak Ridge, Tennessee. pp 325

Cheng, R. 1977. The generation of gamma variables with non-integral shape parameters. *Applied Statistics*. 26: 71-82.

Greenwood, J. A. and Durand, D. 1960. Aids for fitting Gamma distribution by maximum likelihood. *Technometrics*, 2: 55-65

Haan, C. T. 1977. Statistical Methods in Hydrology. IOWA state University Press, Ames, Iowa. pp 378

Nair, B and Sudarsan, K. 1997. Probability analysis of rainfall for crop planning in Jungadh. Unpublished Thesis presented for partial fulfillment of B.Tech., Gujarat Agricultural University, Junagadh. pp 135

Phein, H. N. and Ruksaslip, W. R. 1981. A review of single site models for monthly stream flow generation. *J. Hydrology,* 42: 1-14.

Roldan, R. and Woolhiser, D. 1982. Stochastic daily precipitation models: 1 A comparison of occurrence process. *Water Resources Res.,* 18: 1451-1458.

Shook, R. C. and Highland, H. J. 1969. Probability models with business applications. Richard D Irwin Incorporation, pp 529.

Srikanthan, R. and MacMohan, T. A. 1983. Stochastic generation of annual, monthly, daily evaporation data for Australia. University of Melbourne, Agricultural Engineering Report 63.

Stern, R. D. and Coe R A. 1984. Fitting analysis of daily rainfall data. *J. Roy. Stat. Soc.,* 147:1-18.

Thom, H. C. S. 1958. A note on the gamma distribution. *Mon. Weath. Rev.,* 86 (4):11-13

Tskarais, G., Agrafiotis, G. and Kiountouzis, E. 1984. Modeling the occurrence of wet and dry periods. Proceedings of 5th International Conference on Water Resources Planning and Management, Athens, Grece, pp 5.121-5.130.

Waymire, E. and Gupta, V. K. 1981. The mathematical structure of rainfall representation: a review of stochastic rainfall models. *Water Resources Res.,* 17: 1261-1267.