International Research Journal of Mathematics, Engineering and IT
Vol. 3, Issue 4, April 2016 IF- 3.563 ISSN: (2349-0322)
© Associated Asia Research Foundation (AARF)
Website: www.aarf.asia Email : editor@aarf.asia , editoraarf@gmail.com

A STUDY OF MONSOON DAILY RAINFALL USING INFORMATION THEORY APPROACH

K. K. Saxena & Gurjeet Kaur
Professor of Statistics, Department of Statistics,
University of Dodoma, Dodoma, Tanzania

ABSTRACT

By using the knowledge of the Shannon's entropy, the uncertainty of the transitional system of the south-west monsoon daily rainfall characteristics (such as non- rainy day, light rains, moderate rains, heavy rains, very heavy rains) have been studied in this paper. The favorable or unfavorable condition for different states of Markov chain model of order one during the south-west monsoon period have also been studied by the redundancy test.

Keywords: *Markov chain, redundancy, measure of entropy, uncertainty, redundancy test, transitional probabilities, maximum likelihood estimation.*

Introduction

The agricultural scenario of India is closely linked with the distribution of rainfall throughout the year. Particularly in semi-arid zones rainfall is one of the governing factor in planning agricultural operations. As the distribution of rainfall varies over space and time, it is essential to analyze statistically rainfall data covering long periods for better planning of cropping pattern in semiarid zones, where irrigation facilities are very less and farming is mostly dependent on rainfall

Several researchers have used various statistical technique for studying the behavior of rainfall and related aspects. Fisher (1924) was the first to make a systematic study in this field by studying the influence of rainfall on the yield of wheat at Rothamsted Agricultural Experiment Station and showed that distribution of rainfall during a season is more effective force than its total amount in the yield of a crop. Several authors like Gabriel and Neumann (1962) and Medhi

(1976) have studied the probability of occurrences of dry and wet days through Markovian model. Basu (1988) has studied the probability of occurrences of different rainfall characteristics through one step transitional probability matrices during monsoon period at Maithon by Markovian model. Sarkar and Subramanyam (1995) have studied the characteristics of persisitence and recurrence of precipitation and derived the order of Markov chain. Singh et al.(2002) studied the analysis of statistical behaviour of daily maximum and monthly rainfall. Maruyama et al.(2005) studied Entropy based assessment and clustering of potential water resources availability. Kwarteng et al. (2009) analysed rainfall data for 27 years by using Mann-Kendall test statistics to detect trends in time series data of rainfall. Pechlivanidis et al.(2010) used Entropy for Model diagnostic in rainfall modelling. Saxena et al. (2013) studied the order identification of daily rainfall by using Akaike's Criterion. But Shannon's entropy has not been yet used to study the uncertainties of rainfall and an attempt has been made in this paper to analyze the daily rainfall data of Hisar station for the period from 1925 to 2013 using Shanon's entropy.

Haryana state receives rainfall in the range of 300mm-1200mm in different agro-climatic zones. Hisar is situated at latitude $29^0 10^/$N longitude $75^0 76^/$ E. Its height is 215.2 mt above mean sea level. Hisar receives an average of 318mm rainfall in S-W monsoon which is about 81% of the average annual rainfall. In this study the daily rainfall at Hisar have been categorized into five classes according to definition such as non-rainy days, light rains, moderate rains, heavy rains, very heavy rains. The main objective of this study is to estimate the transitional probabilities of all states of occurrence and to test the uncertainty of transitional probabilities of all state of occurrence by using Shannon's entropy. Data of daily rainfall at Hisar during monsoon period for 89 years (1925-2013) have been used here.

Methodology

Estimation of transitional probabilities: A one step 5X5 transitional probability matrix from one transitional state to another has been found for each of the rainy month from June to September from the frequency of occurrence of daily rainfall amount at Hisar. The Maximum Likelihood method of estimation of transition probabilities daily rainfall has been used. The dis-orderness of such transitional probability is tested from the redundancy of the system. The test of uncertainty of transitional probabilities of all states of occurrence for determining the favourable or unfavourable of the system has been developed through redundancy method.

One step 5X5 transitional probability matrices from one transitional state to another have been computed for each monsoon month. The transitional probability matrix is given by

$$P_{ij} = \begin{pmatrix} P_{00}\ldots\ldots\ldots\ldots P_{04} \\ P_{10}\ldots\ldots\ldots\ldots P_{14} \\ : \qquad\qquad\qquad : \\ : \qquad\qquad\qquad : \\ P_{40}\ldots\ldots\ldots\ldots P_{44} \end{pmatrix}$$

where P_{ij} *(i,j = 0,1,--------4)* is the transitional probability from the j[th] state of occurrence to the i[th] state and $\sum_j P_{ij} = 1$ for each *i*.

Here '0' indicates non rainy days; '1' for light rains, '2' for moderate rain; '3' for heavy rains; '4' for very heavy rain.

Measure of Entropy

A measure of uncertainty (entropy) of the transitional probability (P_{ij}) of the above system is given by Shannon's formula as

$$H_i = \sum_{j=0}^{4} P_{ij} \log P_{ij} \ , \text{ for each } i \ ,$$

where H_i denotes the entropy of the i^{th} state of occurrence (i = 0,1,-----,4).

The entropy H of the Markovian system of transitional probability matrix P_{ij} obtained from the probability of individual state of occurrence P_i and weighted entropy H_i is obtained as under:

$$H = \left\{ -\sum_{i=0}^{4} P_i H_i \right\}, \quad \text{where } P_i \text{ is the probability of the i}^{th} \text{ state of occurrence.}$$

The measure of the entropy *M* of the Markovian model is obtained as

$$M = \left\{ -\sum_i P_i \log P_i \right\} - H$$

The favorable & un-favorableness of the system is tested by redundancy R which is defined by $R = 1 - \left\{ \frac{H}{H_{max}} \right\}$, where H_{max} is the maximum possible entropy and its value is equal to *log*

5= 0.6990. When the redundancy value R tends to 1, the system tends to maximum favorable conditions.

Result and Discussion

The rainfall for south-west monsoon period at Hisar has been categorized into five states according to amount of rainfall received.

Categories	Daily rainfall amount (mm)
Non-rainy days (*D*)	0-6
Light rain (W_1)	6-15
Moderate rain (W_2)	15-35
Heavy rain (W_3)	35-75
Very heavy rain (W_4)	≥75

The conditional transitional probabilities P_{ij} are calculated by dividing the frequencies of one state of occurrence for the day followed from another state of occurrence of the previous day *(fij)* by the frequencies of the state *(fi)*, i.e.

$P_{ij} = f_{ij} / f_i$ and are arranged in the form of a matrix for each monsoon month as shown in Table-1.

Table-1:Transitional probability matrixes for each monsoon month (June-September) at Hisar.

June

$$P_{ij}= \begin{pmatrix} .94 & .04 & .02 & .00 & .00 \\ .75 & .12 & .12 & .01 & .00 \\ .81 & .12 & .05 & .02 & .00 \\ .75 & .08 & .08 & .08 & .00 \\ 1 & .00 & .00 & .00 & .00 \end{pmatrix}$$

July

$$P_{ij}= \begin{pmatrix} .88 & .06 & .04 & .02 & .00 \\ .74 & .09 & .09 & .07 & .01 \\ .71 & .11 & .09 & .08 & .01 \\ .56 & .16 & .18 & .08 & .01 \\ .85 & .15 & .00 & .00 & .00 \end{pmatrix}$$

August

$$P_{ij}= \begin{pmatrix} .89 & .05 & .04 & .02 & .00 \\ .79 & .08 & .12 & .01 & .00 \\ .75 & .08 & .10 & .06 & .00 \\ .68 & .09 & .12 & .11 & .00 \\ .43 & .29 & .14 & .14 & .00 \end{pmatrix}$$

September

$$P_{ij}= \begin{pmatrix} .94 & .03 & .02 & .01 & .00 \\ .79 & .01 & .05 & .04 & .02 \\ .73 & .15 & .05 & .07 & .00 \\ .63 & .06 & .20 & .05 & .06 \\ .33 & .00 & .17 & .50 & .00 \end{pmatrix}$$

Probability of occurrences of all the five states of Markov chain have been calculated and are presented in table-2 for each of the months from June to September.

Table-2: Probability and corresponding information for different individual state of occurrence for each monsoon month (June to September) at Hisar.

	Prob. of Occurrences P_i			
	June	July	August	September
Non-rainy-day P_0	.93	.84	.86	.92
Light rain P_1	.04	.07	.06	.04
Moderate rain P_2	.02	.05	.05	.03
Heavy rain P_3	.01	.03	.03	.01
Very heavy rain P_4	.00	.01	.00	.00

From these probabilities of occerences of various states given in table-2, we now present the Information in bits and [- $\sum P_i \log P_i$] in table-3.

Table-3: Information in bits for each state

	Information in bits $I(P_i) = \log P_i$			
	June	July	August	September
Non-rainyday	-.03	-.08	-.07	-.04
Light rain	-1.39	-1.15	-1.22	-1.39
Moderate rain	-1.69	-1.30	-1.30	-1.52
Heavy rain	-2.0	-1.52	-1.52	-2.0
Very heavy rain	-	-2.0	-	-
- $\sum P_i \log P_i$	0.14	0.28	0.24	0.15

By using the values obtained in Table-3, the weighted entropy values for different categorized states of rainfall amounts for each of the months are calculated. Finally the Redundancy value of R for each month has been calculated. These values are presented in Table-4.

Table-4: Measure of entropy and redundancy at Hisar

	June	July	August	September
Non-rainy days *(H_0)*	0.12	0.21	0.24	0.12
Light rain *(H_1)*	0.33	0.39	0.30	0.35
Moderate rain *(H_2)*	0.28	0.41	0.37	0.37
Heavy rain *(H_3)*	0.36	0.51	0.42	0.48
Very Heavy rain *(H_4)*	0.00	0.18	0.55	0.44
$H = \left\{ -\sum_{i=0}^{4} P_i H_i \right\}$	0.13	0.24	0.21	0.14
Redundancy, R(%)	81	66	70	80
Measure of entropy $M= (-\Sigma P_i \log P_i)- H$	.01	0.04	0.03	0.02

It is clear from Table-4 that the value of the weighted entropy is less during the beginning month June and ending month September as compared to middle months July and august at Hisar.

The redundancy values of the system during monsoon period at Hisar in Table-4 show that the maximum value is 81% during the month of June and minimum value is 66% during July at Hisar. This indicates that favorableness of the Markovian system is maximum in July at Hisar during the monsoon period.

Conclusion

The transitional probability at all states of occurrence for each of the monsoon month at Hisar is estimated by the method of Maximum Likelihood. The entropy (disorderness) of such probability for each of the categorized classes of rainfall amount at Hisar during monsoon period with their weighted entropy values have been determined and the monthly redundancy of the system for each of the monsoon month have been obtained on the basis of weighted entropy and H_{max}.

The results in Table-4 show that the redundancy value is more during the month of June and September as compare to its value in July and August. Since this redundancy value is used to determine the favorableness or unfavorablness of the system, as the redundancy value R tends to 1, the system tends to maximum favorable conditions i.e. almost certain. Hence, at Hisar during S-W monsoon season, the Markovian system has been found to be favorable by redundancy test.

References

Basu, G.C.(1988): A study of monsoon daily rainfall at Maithon by markovian model and information theory. *Mausam.* 39(1): 83-86

Fisher, R.A.(1924): Influence of rainfall on the yield of wheat at Rathamsted. *Phel. Trans. Roy. Stat.* London. B 213: 89-142.

Gabriel, K.R. and Neumann,J.(1962): A Markov chain model for daily rainfall in *Tel- Aviv. Quart. J.R. Met. Soc.* 88: 90-95.

Kwarteng,A.Y., Dorvlo,A.S. and Kumar,G.T.V. (2009): Analysis of a 27-year rainfall data (1977-2003) in the Sultanate of Oman. *International Jr. of Climatology,29:605-617.*

Maruyama,T., Kawachi,T and Singh,V.P. (2005): Entropy based assessment and clustering of potential water resources availability. *Jr. of Hydrology.*309 (1-4): 104-113.

Penchivanidis, I.G., Jackson,B, and McMillan,H. (2010): In the proceedings of *International Environmental Modelling and Software Society. (http://www.iemss.org/iemss2010/index=Main proceedings).*

Medhi, J.(1976). A Markov chain model for the occurrence of dry and wet days. *Ind. J. Met. Hyd. and Geophys.* 27(4): 431-435.

Sarkar, J. and Subramanyam,S.K. (1995): Dry spell probability by Markov chain model at Madras during the north east monsoon months. *Mausam.*46(1),85-87.

Singh, C.V., Adhikari,R.S. and Garg,H.P. (2002):Analysis of the statistical behaviour of daily maximum amd monthly rainfall data at new delhi during monsoon period.*Advances in Atmospheric Sciences.* Vol-19,3. 425-432.

Saxena, K.K., ruhal, Lokesh and Rao,V.U.M. (2013): Akaike Information Criterion for order identification of Daily rainfall at Hisar. *International Jr. Agricult.Stat.Sci* . Vol 9,1.71-78.