Universal Journal of Environmental Research and Technology
 eISSN 2249 0256
Available Online at: www.environmentaljournal.org
Volume 1, Issue 4: 566-570

Open Access **Short Communication**

Stochastic Modelling of Annual Rainfall at Tamil Nadu

Tamil Selvi S., Samuel Selvaraj R.*

Department of Physics, Shree Motilal Kanhaiyalal Fomra Institute Of Technology, Chennai.
*Department Of Physics, Presidency College, Chennai

Corresponding author: tssekar9773@yahoo.com, samuelselvaraj.r@rediffmail.com

Abstract:
Rainfall is a phenomenon, which directly or indirectly affects all the sectors like agriculture, insurance, industry and other allied fields. Prediction of rainfall has remained an unsolved problem till now. One of the statistical techniques is the Markov chain used to predict precipitation on short term, at meteorological stations. This paper deals with the variations of annual rainfall in Tamil Nadu based on Markov Chain models. For this purpose, we derived annual rainfall from 1901 to 2000 and frequency distribution table is formed. To calculate the yearly rainfall variations the class interval is treated as states and transition probability matrix is formed. The uniform random states are also formed by generating uniform random number. If the future climate conditions are known with sufficient accuracy, the stochastic climate models available at present can be adapted to generate climate for the new conditions.

Keywords: Markov chain, frequency distribution, random number, rainfall, Prediction

1.0 Introduction:
The pattern and amount of rainfall are among the most important factors that affect agricultural systems. The analysis of rainfall records for long periods provides information about rainfall patterns and variability (Lazaro, R. et.al; 2001). Rain plays a major role in hydrology that finds its greatest applications in the design and operations of water resources, engineering works as well as agricultural systems (Srikanthan, R. and McMahon, T. A.,2005a). There are four techniques used in rainfall prediction. Statistical technique, Stochastic method, Artificial Neural Network and Numerical weather prediction. There are primarily three methods of weather forecasting such as Synoptic, Statistical and Numerical methods. Rainmann model and Lorentz model are also used for forecasting (Kelkar, R.R, 2009). An application of stochastic process for describing and analyzing the annual rainfall pattern at Tamil Nadu is presented. A model based on the first-order Markov chain was developed. The probability estimation of rainfall states from available time series helps to obtain predictions for rainfall statistical parameters such as the averages, standard deviations and the coefficient of variation. Most of the models used in the past do not take into account the year to year variations in the model parameters. They were assumed to be constant from year to year and only the within-year seasonal variations in parameters were taken into account. Recently, (Thyer and Kuczera 1999, 2000) developed a hidden state Markov model to account the long term persistence in annual rainfall. The main purpose of this paper is to show the use of first order Markov Chain modeling for yearly basis of rainfall measurements over Tamil Nadu. The annual rainfall from 1901 to 2000 is used and the frequency distribution table is formed. The class interval is treated as states and then uncertainty under various states occupied by formation of transition probability matrix.

2.0 Methodology:
Tamil Nadu is the southernmost state of India, surrounded by Andhra Pradesh from the North, Karnataka and Kerala from the west, Indian Ocean from the south and Bay of Bengal from the East. Tamil Nadu roughly extends between the 8° 04' N latitude (Cape Comorin) and the 78° 0' E longitude. Geographically, Tamil Nadu is situated on the eastern side of the Indian Peninsula between the northern latitude of 8.5" and 13.35" and the eastern longitude of 76.15" and 80.20". The average annual rainfall in Tamil Nadu ranges between 25 and 75 inches (635 and 1,905 mm) a year.

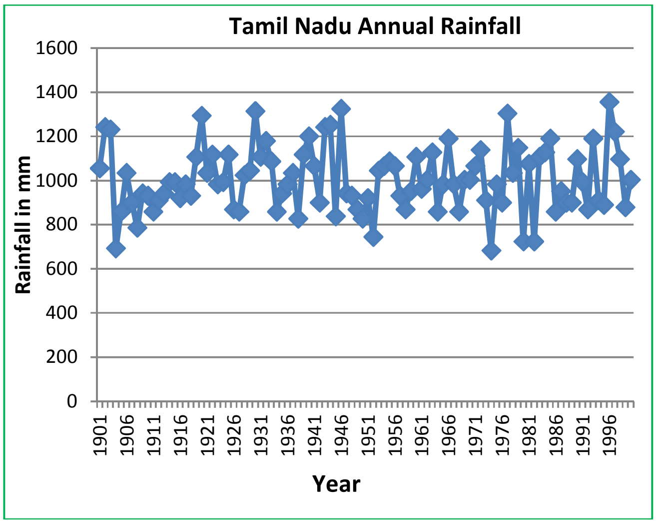

Figure a: Fluctuations of annual rainfall at Tamil Nadu

The data used in this study were obtained from Regional Meteorological Department, Chennai from the year 1901 to 2000. From the fluctuations of extreme yearly rainfall depicted in Fig.-a time series is prepared and converted into rainfall states by preparing suitable frequency distribution table. In the frequency distribution table (Table-1), each class supposed as specific states of rainfall and denoted by capital alphabet A, B, C, D, E, F and G. Table-2 revealed that during 100 years the average rainfall was 1005.583 mm/year and the coefficient of variation was observed nearly 14%. The positive kurtosis indicates heavy tails and peakedness relative to the normal distribution, whereas negative kurtosis indicates light tails and flatness.

Table 1: Frequency distribution table for annual rainfall

Class Interval	States	Frequency
680.1 - 779.9	A	5
779.9 - 879.9	B	15
879.9 - 979.9	C	23
979.9 - 1079.9	D	27
1079.9 - 1179.9	E	12
1179.9 - 1279.9	F	13
1279.9 - 1379.9	G	5

Table-2: Summary statistics of annual rainfall data

Statistics	Value
Sample size	100
Range	1356.4
Mean	1005.583
Variance	21283.34
Standard Deviation	145.8881
Coefficient of Variation	14.5078
Skewness	0.244383
Kurtosis	-0.2186

2.1 Markov-Chain Modeling:

A Russian mathematician, Markov, introduced the concept of a process (later named after him 'a Markov process') in which a sequence or chain of discrete states in time for which the probability of transition from one state to any given state in the next step in the chain depends on the condition during the previous step (Zohadie Bardaie, M. and Ahmad Che Abdul Salam., 1981). (Gabriel and Neumann,1962) analyzed the occurrence of rain by fitting a two-state, first-order Markov chain. (Carey and Haan, 1978) used multi-state Markov chain models to generate daily rainfall depths. Rainfall exhibits a strong variability in time and space. Hence its stochastic modeling is not an easy task (De Michele and Bernardara 2005). The Markov chain models have two advantages: (1) the forecasts are available immediately after the observations are done because the use as predictors only the local information on the weather and (2) they need minimal computation after the climatological data have been processed. A first-order Markov chain is one in which knowing one variable (like cloudiness, precipitation amount, temperature, fog, frost, wind) at time t is sufficient to forecast it at some later time. A first order Markov chain is a stochastic process having the property that the value of the process at time t, X_t , depends only on its value at time t-1, X_{t-1}, and not on the sequence of values that the process passed through in arriving at X_{t-1}. A 'C state' Markov chain requires that C(C-1) transition probabilities be estimated and the remaining C Pij can be determined using the relation

$$\sum_{j=1} \boldsymbol{pij} = \mathbf{1} \qquad \textbf{(1)}$$

The C^2 transition probabilities are given by the stochastic matrix P.

$$\mathbf{P} = \begin{pmatrix} p11 & p12 & \dots & p1c \\ p21 & p22 & \dots & p2c \\ \dots & \dots & \dots & \dots \\ pc1 & pc2 & \dots & pcc \end{pmatrix} \qquad \textbf{(2)}$$

Once P is known, all that is required to determine the probabilistic behavior of the Markov Chain is the initial state of the chain. In the following, $pj^{(n\ )}$ denotes the probability that the chain is in state j at step or time n. The 1xC vector $p^{(n)}$ has elements $p_j^{(n)}$. Thus

$$\mathbf{p^{(n)} = [\ p_1^{(n)}, p_2^{(n)}, \dots, p_c^{(n)}\ ] \text{ and } p^{(1)} = p^{(0)}\ p} \qquad \textbf{(3)}$$

Where, $p^{(0)}$ is the initial probability vector.

In general, $\mathbf{p^{(n+c)} = p^{(c)}\ p^{(n)}}$ **(4)**

Where, p^n is the n th power of p.

2.2 Parameter Estimation:

The parameters for the occurrence model are transition probabilities, p_{ij} s, which forms the transition matrix P. The estimate for p_{ij} is given by

$$P_{ij} = \frac{\mathbf{n_{ij}}}{\sum_{j=1}^{c} nij}$$

where, n_{ij} is the number of times the observed data went from state i to state j. In order to calculate the Markov chain transitional probabilities, initially the rainfall variation domain is divided into many states according to the frequency distribution table.

3.0 Results and Discussion:

In the following matrix, the annual rainfall is presented with seven states in the form of population probability transition matrix as

$$\mathbf{p_{ij}} = \begin{pmatrix} \frac{0}{5} & \frac{1}{5} & \frac{0}{5} & \frac{3}{5} & \frac{1}{5} & \frac{0}{5} & \frac{0}{5} \\ \frac{0}{16} & \frac{2}{16} & \frac{7}{16} & \frac{4}{16} & \frac{0}{16} & \frac{2}{16} & \frac{1}{16} \\ \frac{2}{24} & \frac{4}{24} & \frac{7}{24} & \frac{5}{24} & \frac{3}{24} & \frac{1}{24} & \frac{2}{24} \\ \frac{1}{26} & \frac{3}{26} & \frac{6}{26} & \frac{7}{26} & \frac{4}{26} & \frac{4}{26} & \frac{1}{26} \\ \frac{1}{11} & \frac{2}{11} & \frac{2}{11} & \frac{2}{11} & \frac{1}{11} & \frac{2}{11} & \frac{1}{11} \\ \frac{1}{13} & \frac{3}{13} & \frac{1}{13} & \frac{3}{13} & \frac{2}{13} & \frac{3}{13} & \frac{0}{13} \\ \frac{0}{5} & \frac{0}{5} & \frac{1}{5} & \frac{2}{5} & \frac{1}{5} & \frac{1}{5} & \frac{0}{5} \end{pmatrix}$$

$$
= \begin{pmatrix}
0.000 & 0.200 & 0.000 & 0.600 & 0.200 & 0.000 & 0.000 \\
0.000 & 0.125 & 0.438 & 0.250 & 0.000 & 0.125 & 0.063 \\
0.083 & 0.167 & 0.292 & 0.208 & 0.125 & 0.042 & 0.083 \\
0.038 & 0.115 & 0.231 & 0.269 & 0.154 & 0.154 & 0.038 \\
0.091 & 0.182 & 0.182 & 0.182 & 0.091 & 0.182 & 0.091 \\
0.077 & 0.231 & 0.077 & 0.231 & 0.154 & 0.231 & 0.000 \\
0.000 & 0.000 & 0.200 & 0.400 & 0.200 & 0.200 & 0.000
\end{pmatrix}
$$

This matrix provides the basis of future likely synthetic rainfall states generations. The following steps are necessary for the generation of extreme rainfall states.

1) Calculate the cumulative probability transition matrix, Pc in which each row ends with 1. Hence, cumulative summation within each row leads to

$$
Pc = \begin{pmatrix}
0.000 & 0.200 & 0.200 & 0.800 & 1.000 & 1.000 & 1.000 \\
0.000 & 0.125 & 0.563 & 0.813 & 0.813 & 0.938 & 1.000 \\
0.083 & 0.250 & 0.542 & 0.750 & 0.875 & 0.917 & 1.000 \\
0.038 & 0.153 & 0.384 & 0.653 & 0.807 & 0.961 & 1.000 \\
0.091 & 0.273 & 0.455 & 0.637 & 0.728 & 0.910 & 1.000 \\
0.077 & 0.308 & 0.385 & 0.616 & 0.770 & 1.000 & 1.000 \\
0.000 & 0.000 & 0.200 & 0.600 & 0.800 & 1.000 & 1.000
\end{pmatrix}
$$

2) The primary state is adopted randomly and then by using a uniform random number generator, a random value is generated within the range of 0 and 1.

2) The next state is obtained where this random value is greater than the cumulative probability of the previous state but less than or equal to the cumulative probability of the following state. In this way, any desired number of rainfall states can be generated. Here in this study the states A, B, C, D, E, F and G are the states and they are the class interval value as it appeared in Table- 1, so when we generate the synthetic series for these states we can get the states of rainfall not the exact value (amount in mm) of rainfall. The states will be decomposed according to the following rule,

$$
\textbf{State} = \begin{cases}
A \;\; if \;\; u < 0.083 \\
B \;\; if \;\; 0.083 \le u \le 0.25 \\
C \;\; if \;\; 0.25 \le u \le 0.542 \\
D \;\; if \;\; 0.542 \le u \le 0.75 \\
E \;\; if \;\; 0.75 \le u \le 0.875 \\
F \;\; if \;\; 0.875 \le u \le 0.917 \\
G \;\; if \;\; 0.917 \le u \le 1
\end{cases}
$$

Where u is the generated uniform random number.

The Table-3 is clearly shown how the states change with changing the probability where each state (English alphabets) is attached with some class interval. This type of simulation is very needful to generate interval of rainfall when our instrument malfunctioned during some short period of time. It is also beneficial to Aviation Company to make short term probabilistic prediction for climatic conditions. The variability of rainfall and the pattern of extreme high or low precipitation are very important for the agriculture as well as the economy of the country. It is well established that the rainfall is changing on both the global and the regional scales (Hulme *et al*. 1998, Dore 2005, Kayano and Sans´ıgolo, 2008) due to global warming. Concerns over climate change caused by increasing concentration of CO2 and other trace gases in the atmosphere has increased in recent years. A major effect of climate change may be alterations in regional hydrologic cycles and changes in regional water availability. The use of modified water balance models offers many advantages in evaluating the regional impacts of global climate change (Gleick, 1989).

Table 3: Generation of synthetic series of Rainfall Uniform Random Number

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Uniform Random Number	**0.139**	**0.952**	**0.278**	**0.941**	**0.165**	**0.660**	**0.837**	**0.106**	**0.420**	**0.473**	**0.170**	**0.236**
Random States	**B**	**G**	**C**	**G**	**B**	**D**	**E**	**B**	**C**	**C**	**B**	**B**

4.0 Conclusion:

The transition probability matrix represents the weather model in which the trend of the following year is estimated. But the long range forecasting based on this model does not give more accuracy. The year 2001 and 2002 actual rainfall values do not match with our model. Although the model described here is used only to forecast the future trend, the same Markov chain can be further developed to forecast with high accuracy. The analysis of extreme yearly rainfall shows that Markov Chain approach provides one alternative of modelling future variation in rainfall. These variations may either be in the form too much water which will lead to flooding or too little water, which will lead to draught. Markov modelling is one of the tools that can be utilized to assist planners in assessing the rainfall.

References:

1) Carey, D.I. and Haan, C.T., (1978), Markov process for simulating daily point rainfall*, J.Irrig. and Drain Div.,* 104, 111 – 125.
2) De Michele, C. and Bernardara, P. (2005). Spectral analysis and modeling of space-time *Atmosph. Res*., 124–136.
3) Dore, MHI. (2005). Climate change and changes in global precipitation patterns: what do we know. *Environ. In*. 31(8): 1167-1181.
4) Gabriel,K.R. and Neumann, J., (1962), A Markov chain model for daily rainfall occurrences at Tel Aviv, Quar.J. Roy. Meteor.Soc., 88, 90 – 95.
5) Gleick, P. H. 1989. Climate change, hydrology, and water resources. *Reviews of Geophysics*, 27: 329-344.
6) Kayano, M. T. and Sans´ıgolo, C. (2008). Interannual to decadal variations of precipitation and daily maximum and daily minimum temperatures in southern Brazil. *Theoretical and App. Clim.* 97(1-2): 81-90.
7) Hulme, M., Osborn, T. J. and Johns, T. C. (1998). Precipitation sensitivity to global warming: comparison of observations with HADCM2 simulations.*Geophysical Res. Letter* 25: 3379-3382.
8) Kelkar, R.R, 2009 , Monsoon Prediction, (B.S. Publications, Hyderabad, India)
9) Lazaro, R., Rodrigo, FS., Gutierrez, L., Puigdefafregas, J., (2001), Analysis of a 30-year rainfall record (1967-1997) in semi-arid SE Spain for implications on vegetation, J. Arid Environ., 48, 373 - 395.
10) Srikanthan, R. and McMahon, T. A. (2005a); Automatic Evaluation of Stochastically Generated Rainfall Data**,** *Conference Paper, Australian Journal of Water Resources 2005,* Vol. 8, No. 2, 195-201.
11) Thyer, M. and Kuzcera, G. (1999). Modelling long-term persistence in rainfall time series: Sydney rainfall case study. Hydrology and Water Resources Symposium, Institution of Engineer, Australia: 550-555.
12) Thyer, M. and Kuzcera, G. (2000). Modelling long-term persistence in hydro-climatic time series using a hidden state Markov model. Submitted to Water Resources Research
13) Zohadie Bardaie, M. and Ahmad Che Abdul Salam., (1981), A Stochastic Model of Daily Rainfall for Universiti Pertanian Malaysia, Serdang, Pertanika, 4(1), 1 - 9