International Journal of Scientific Engineering and Technology ISSN:2277-1581
Volume No.6 Issue No.4, pp: 128-131 01 April.2017

Rainfall Forecasting in Gasabo District Using Markov Chain Properties

Constantin Tuyizere[1], Joseph Kyalo Mung'atu[1], Denis Ndanguza[2]
[1]Jomo Kenyatta University of Agriculture and Technology-Kigali Campus, Department of Applied Statistics, P.O. Box 3373, Kigali- Rwanda
[2]University of Rwanda, College of Science and Technology, P.O.Box 3300 Kigali-Rwanda

Abstract: Over the last decade, the whole world has undergone the atmosphere changes; as a result, the disasters such as floods, landslides, droughts and ground water declination have taken place. The same problem has remarkably affected the developing countries including Rwanda, due to the weakness of rainfall forecast using meteorological historical data. The researcher developed a rainfall Markov chain model based on Rwandan meteorological data collected for 44 years (1971-2014). The findings from survey indicated the transition matrix and limiting states probabilities vector as $\pi=[0.78\ \ 0.07\ \ 0.12\ \ 0.03], \Rightarrow \sum \pi_i = 1,$ ***got after 6 years. The correlation between states has been found, and revealed that there were other uncontrollable factors. The researcher recommends decision makers of Rwanda Meteorology Agency to adopt the markov chain model to provide and to disseminate the needed information about climate change early and do the forecasting.***

Keywords: Stochastic process, Markov chain model, transition probability matrix, limiting state probability vector.

1. Introduction

Recently, many researchers in climatology area have revealed the changes in the normal atmospheric conditions and weather patterns. They show that the key environmental problems are land degradation, deforestation, wetland and biodiversity loss [1]. Climate change is expected to increase vulnerability to existing stresses mentioned above and these uncertain changes of weather cause disasters such as floods, landslides, drought, ground water declination, which affect life of human beings.

Gasabo is a district of Rwanda where there is ostensibly the vulnerability in terms of climate change. Thus, it affects both the environment and ecosystems. Since the end of 2015, there have been innumerable deficits. Furthermore, the floods, landslides, and heavy rainfall have caused a considerable toll throughout the country. Due to such problem, every individual, family or organization has to prevent the said catastrophe. The results of the present research suggest that the solution of the abovementioned problem be forecasted in the long term using Markov chain model.

Accurate and timely rainfall prediction is a major challenge for the scientific community. Several recent research studies have developed rainfall prediction using different weather and climate forecasting methods.

Gabriel and Neumann [2] studied the sequence of daily rainfall occurrence. They found that the daily rainfall occurrence for Tel Aviv data was successfully fitted with the first-order Markov chain model.

Abubakar Usman Yusuf et. al [3] demonstrate the application of Markov chain model to study the annual rainfall data in Minna in the North Central geo-political zone of Nigeria with respect to crop production. In view of the uncertainty of annual rainfall and crop production in the region, the annual rainfall of the present year could be used to make a forecast for the following year(s) and in the long run.

Tamil and Selvaraj [4] used Markov chain model to calculate the yearly rainfall variations such that the class interval is treated as states. They found that 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.

Recently by Raheem M.A. et. al [5] a three-state Markov chain was employed to examine the pattern and distribution of daily rainfall in Uyo metropolis of Nigeria. The expected length of spell of each of the three seasons (dry, wet and rainy) and weather cycle for each of the periods of pre-monsoon, monsoon and post-monsoon were all determined for Uyo metropolis using the rainfall data collected.

In this paper, the researcher is interested to develop a rainfall Markov chain model based on Rwandan meteorological data collected for 44 years (1971-2014), to overcome the weakness of forecasting in long run term.

2. Theoretical review on Markov chain

According to William in [6], a discrete-time Markov chain $\{X_n, n = 0, 1, 2, \ldots\}$ is a stochastic process that satisfies the following relationship, called the markov property:
For all natural numbers n and all states x_n,

$$\text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0\} = \text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n\}. \quad \textbf{(2.1)}$$

To simplify the notation, the conditional probabilities

International Journal of Scientific Engineering and Technology ISSN:2277-1581
Volume No.6 Issue No.4, pp: 128-131 01 April.2017

Prob $\{X_{n+1} = x_{n+1} | X_n = x_n\}$, now written as Prob $\{X_{n+1} = j | X_n = i\}$, are called the single-step transition probabilities, or just the transition probabilities, of the Markov chain. They give the conditional probability of making a transition from state $x_n = i$ to state $x_{n+1} = j$ when the time parameter increases from n to n + 1. They are denoted by

$$p_{ij}(n) = \Pr ob\{X_{n+1} = j | X_n = i\} \quad \textbf{(2.2)}$$

The matrix $P(n)$, formed by placing $p_{ij}(n)$ in row i and column j, for all i and j, is called the transition probability matrix or chain matrix. We have

$$P(n) = \begin{bmatrix} p_{oo}(n) & p_{01}(n) & p_{02}(n) & \cdots & p_{0j}(n) & \cdots \\ p_{10}(n) & p_{11}(n) & p_{12}(n) & \cdots & p_{1j}(n) & \cdots \\ p_{20}(n) & p_{21}(n) & p_{22}(n) & \cdots & p_{2j}(n) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{i0}(n) & p_{i1}(n) & p_{i2}(n) & \cdots & p_{ij}(n) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Notice that the elements of the matrix $P(n)$ satisfy the following two properties: $0 \le p_{ij}(n) \le 1$,

and, for all i, $\sum_j p_{ij}(n) = 1$, only sum of elements on each row.

A matrix that satisfies these properties is called a Markov matrix or stochastic matrix.

2.1 The Chapman-Kolmogorov Equations

We have already defined the one step transition probabilities p_{ij}. We now define the n-step transition probabilities p_{ij}^n to be the probability that a process in state **i** will be in state **j** after **n** additional transitions.

Let sum over all sample paths of length 2 that lead from state **i** to state **k**. For a homogeneous discrete-time Markov chain, we have, for any n = 0, 1, 2, . . .

Prob $\{X_{n+2} = k, X_{n+1} = j | X_n = i\} = p_{ij} p_{jk}$, and from the theorem of total probability, we obtain

Prob $\{X_{n+2} = k | X_n = i\} = \sum_{all\ j} p_{ij} p_{jk}$.

Observe that the right-hand side defines the ik[th] element of the matrix obtained when P is multiplied with itself, i.e., the element $\left(p^2\right)_{ik} \equiv p_{ik}^{(2)}$. Continuing in this fashion, and observing that

Prob $\{X_{n+3} = l, X_{n+2} = k, X_{n+1} = j | X_n = i\} = p_{ij} p_{jk} p_{kl}$,

we obtain, once again with the aid of the theorem of total probability,

$$\Pr ob\{X_{n+3} = 1 | X_n = i\}$$
$$= \sum_{all\ j} \sum_{all\ k} p_{ij} p_{jk} p_{kl}$$
$$= \sum_{all\ j} p_{ij} \sum_{all\ k} p_{jk} p_{kl}$$
$$= \sum_{all\ j} p_{ij} p_{jl}^{(2)}$$

which is the il[th] element of P^3, $p_{il}^{(2)}$ being the jl element of P^2.

It follows that we may generalize the single-step transition probability matrix of a homogeneous Markov chain to an *m*-step transition probability matrix whose elements $p_{ij}^{(m)}$ = Prob $\{X_{n+m} = j | X_n = i\}$ can be obtained from the single-step transition probabilities.

In matrix notation, the Chapman-Kolmogorov equations are written as $P^{(m)} = P^{(l)} P^{(m-l)}$ where, by definition, $P^{(0)} = I$ the identity matrix.

Then we can write

$$P^{(n+1)} = P^{(n)} P \quad \textbf{(2.3)}$$

Let $\pi_i^{(n)}$ be the probability that the Markov Chain begins in state i, and let $\pi^{(0)}$ be the row vector whose i^{th} element is $\pi_i^{(0)}$.

Then the j^{th} element of the vector that results from forming the product $\pi^{(0)} P(0)$ gives the probability of being in state j after the first time step. We write this as $\pi^{(1)} = \pi^{(0)} P(0)$

.

For a homogeneous Markov Chain, this becomes

$$\pi^{(1)} = \pi^{(0)} P \quad \textbf{(2.4)}$$

The elements of the vector π^1 provide the probability of being in the various states of the Markov Chain (i.e the probability distribution) after the first time step.

To compute the probability of being in any state after two time steps, we need to form

$$\pi^{(2)} = \pi^{(1)} P(1) = \pi^{(0)} P(0) P(1)$$

For a homogeneous Markov Chain, this becomes

$$\pi^{(2)} = \pi^{(1)} P = \pi^{(0)} P^2 \quad \textbf{(2.5)}$$

ISSN:2277-1581
01 April.2017

In general after n steps, the probability distribution is given by: $\pi^{(n)}=\pi^{(n-1)}P(n-1)=\pi^{(0)}P(0)P(1)....P(n-1)$ 0r, for a homogeneous Markov Chain

$$\pi^{(n)}=\pi^{(n-1)}P=\pi^{(0)}P^n \qquad \textbf{(2.6)}$$

2.2 Limiting Probabilities

When iterations are performed, it seems that p_{ij}^{n} is converging to some value, as $n \to \infty$, which is the same for all i. In other words, there exists a limiting probability that the process will be in state j after a large number of transitions, and this value is independent of the initial state.

Consider our case the equilibrium distribution is $\pi=[\pi_1\ \pi_2\ \pi_3\ \pi_4]$.

If we let $n \longrightarrow \infty$ in equation (2.6) we have

$$\pi=\pi P \qquad \textbf{(2.7)}$$

and $\sum_{i=1}^{4}\pi_i=1$

3. Methodology and material

In this paper researchers design the descriptive survey and the entire population is used. Different computer tools, such as Microsoft Excel, SPSS and Matlab are used for analysis of data and interpretation of results. A four-state Markov chain is used to describe the behavior of rainfall occurrences in Gasabo District. The states, as considered are: dry day (d), low rainfall (l), moderate rainfall (m) and high rainfall (h).

The conditions of rainfall occurrence for the four states are defined as follows: a day will be considered dry if rainfall occurrence on that day will be not more than 2.5 mm, low rain if rainfall occurrence ranges from 2.5 mm to less than 5.00 mm, moderate rain if rainfall occurrence ranges from 5.00 mm to less than 20 mm and high rain if rainfall occurrence ranges from 20 mm and above.

The probability of the process being in a particular state, $P_{ij}, i,j=\{d,l,m,h\}$, is calculated based on the Markov chain assumption that attaining a state depends on the immediate preceding state only.

The maximum likelihood estimators of $P_{ij}, i,j=\{d,l,m,h\}$, $\hat{P}_{ij}$ are given by

$$\hat{P}_{ij}=\frac{n_{ij}}{\sum_{j=d}^{r} n_{ij}}. \qquad \textbf{(2.8)}$$

$n_{ij}, i,j=\{d,l,m,h\}$ is the number of **j** days preceded by i days.

Markov chain, transition probability matrix, probabilities of dry day, low rainfall, moderate rainfall and high rainfall in the long run (equilibrium), is determined in the analysis of daily rainfall data.

After finding transition probability matrix and n-step transition probabilities, following HUFTY André [7], we proceed to test the relationship between the current day and previous day as follows:

Let consider (O) the dry day and I the wet day,

$$r = p_{II} - p_{OI} \qquad \textbf{(3.1)}$$

with p_{II} probability of a wet day preceded by a wet day

and p_{OI} probability of a wet day preceded by a dry day.

4. Results and Discussion

In this present paper, the daily rainfall data for 41 years from Jan. 1, 1971 to Dec. 31, 2014 was collected from Kigali station located in Gasabo District. The rainfall occurrence, based on daily rainfall, was studied.

The yearly maximum daily rainfall based on 41 years is presented in Figure 1.

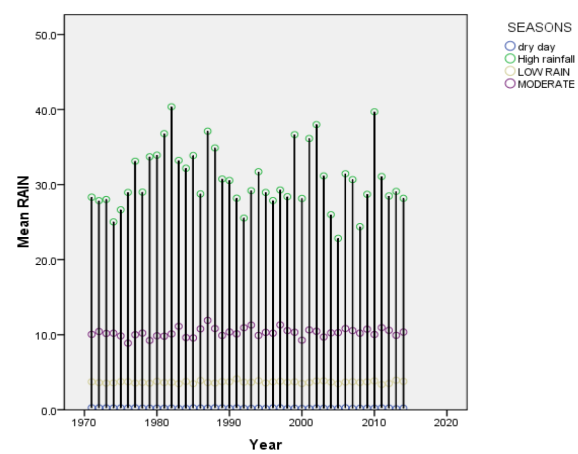

Figure 1: A summary of annual rainfall in Gasabo District between 1971- 2014 and states distribution

From equation (2.8), using the maximum likelihood estimator, we obtain the transition probability matrix P

$$P=\begin{bmatrix}0.813 & 0.056 & 0.103 & 0.028\\ 0.667 & 0.103 & 0.185 & 0.045\\ 0.674 & 0.091 & 0.189 & 0.046\\ 0.618 & 0.102 & 0.216 & 0.064\end{bmatrix}$$

The n-step transition probability

From equation (2.3) after six step we find, on iteration, the transition probability matrix P^6

$$P^6=\begin{bmatrix}0.780 & 0.065 & 0.123 & 0.033\\ 0.780 & 0.065 & 0.123 & 0.033\\ 0.780 & 0.065 & 0.123 & 0.033\\ 0.780 & 0.065 & 0.123 & 0.033\end{bmatrix} \quad \textbf{(4.1)}$$

The limiting state probabilities

Known that the transition probabilities stabilizes to $n \geq 6$, with the initial state probability vector $[0\ \ 0\ \ 1\ \ 0]$ we have

$$\pi^{(0)}P^n=[0\ \ 0\ \ 1\ \ 0]\begin{bmatrix}0.78 & 0.07 & 0.12 & 0.03\\ 0.78 & 0.07 & 0.12 & 0.03\\ 0.78 & 0.07 & 0.12 & 0.03\\ 0.78 & 0.07 & 0.12 & 0.03\end{bmatrix}$$

$$=[0.78\ \ 0.07\ \ 0.12\ \ 0.03]$$

From equation (2.7) and (4.1) the limiting state probability is given by:

$$\pi=\pi p=[0.78\ \ 0.07\ \ 0.12\ \ 0.03]$$

The results show that the probability to have dry, low rainfall, moderate rainfall and high rainfall in the first year, given that it is a moderate at present are 0.674, 0.091, 0.189 and 0.046 respectively. The first increases to 0.78 after six years and the rest drop to 0.07, 0.12 and 0.3 respectively. This means in the long-run 78% of annual rainfall in Gasabo District will be dry, 7% will be low rainfall, and 12% will be moderate rainfall and 3% high annual rainfall. These equilibrium probabilities are independent of any initial state.

5. Conclusions

In this paper we have shown how markov chain model plays a big role to forecast rainfall in the long run. This method facilitates effective mathematical computation to find spatial and temporal changes in rainfall pattern, intensity and concentration of an area, another advantage is that: the forecasts are available immediately after the observations are done because the use as predictors only the local information on the weather. The researcher recommends decision makers of Rwanda Meteorology Agency to adopt the markov chain model to provide and to disseminate the needed information about climate change very soon and do the forecasting of the climate change for the long run.

6. References

i. UNEP, (2011), From post-conflict to environmentally sustainable development

ii. Gabriel, K. R. and Neumann, J. (1962). A markov chain model for daily rainfall occurrences at Tel Aviv. Quart. J. Roy. Met. Soc 88:90-95

iii. Abubakar U.Y., Lawal A.& Muhammed A.(2014). Markov Chain Model and its application to annual rainfall distribution for crop production. American Journal of Theoretical and Applied statistics 3(2): 39-43

iv.Tamil Selvi S. & Samuel Selvaraj R.(2011). Stochastic Modeling of Annual Rainfall at Tamil Nadu. Universal Journal of Environmental Research and Technology. 1(4) 566-570

v. Raheem M.A., Yahya W.B.,& Obisesanat K.O. (2015). A Markov chain approach on pattern of rainfall distribution. Journal of Environmental Statistics.7 (1), 1-13

vi. Stewart W.J. (2009). Probability, Markov chains, Queues, and simulation. Princeton University Press. Oxford

vii.] HUFTY André (2001), Introduction ā la climatologie, les Presses de l'Université Laval, Bruxelles