

Stochastic Modeling of Annual Rainfall for Crop Production in North Central Nigeria

Lawal. A, Abubakar U.Y, D. Hakimi and M.D. Shehu

Department of Mathematics, Federal University of Technology Minna, Nigeria

Abstract

A stochastic model to study annual rainfall pattern in North Central Nigeria has been presented in this paper. A Hidden Markov Model (HMM) was developed for the study. The study classified amount of rainfall at a time into three states, each state with eight possible observations. The HMM was trained using Baum-Welch algorithm to attained maximum likelihood, after which it was used to make predictions. The Model was implemented in Niger, Benue and Plateau states of North Central Nigeria. Results from locations in Niger and Benue states showed some similarities, as compared to location in Plateau state with different pattern. The similarities may suggest one of the reasons that makes both states the leading producers of food crops in the country. The validity test for the model showed that, the model is reliable and dependable. Therefore, results from this model could serve as a guide to the farmers and the government to plan strategies for high crop production in region. The results could also assist the residents in this region to better understand the dynamics of rainfall which may be helpful for effective planning and viable productions.

Keywords: Hidden Markov Model, Annual Rainfall, North Central Nigeria, Baum-Welch Algorithm.

1.0 Introduction

Annual rainfall varies in Nigeria. It decreases from the south to the north, rainfall comes late and not evenly spread across the rainy season and consequently has adverse effect on agricultural activities [1]. In each year, rainfall onset and recession have become a very big problem to the farmers and policy makers in planning for crop cultivation in north central Nigeria. This is because rainfall starts early or late and stops early or late in a year. These variations on rainfall onset, recession, amount and spread had over the years led to improper crop planning and cultivation, and it had consequently led to poor harvest. The majority of the people living in this part of the country are farmers and rainfall is the major source of water for agricultural activities. Because of the dependence of agricultural production on rainfall variability and quantity, and the unpredictable nature of the rainfall in this part of the country, had over the years led to: improper crop planning and cultivation, poor harvest, lost of income of the farmers, shortage of food to the country, hydrological extremes such as mud-flows, floods, landslides, droughts and debris. It has generally reduced Gross Domestic Product (GDP) of the country there by affecting its economy. In view of the above mentioned problems, the research, is aimed at developing a stochastic mathematical model that can predict rainfall onset, recession, amount and spread within a year, this will help provides some quantitative information to the farmers and government that could assists in boosting crop production and also reduce the danger of hydrological extremes.

The research, considers the use of stochastic tool called Hidden Markov Model(HMM). HMMs are extensions of Markov models where each observation is the result of a stochastic process in one of several unobserved states. HMM has been successfully applied in automatic speech recognition and speech synthesis [2], molecular biology for DNA and protein sequencing [3], signal processing [4], bioinformatics [5], telecommunication [6] and in pattern recognition [7].

1.1 The Study Area

The study area of this research is north central Nigeria. However, it is of three states out of the six states from this region. The selected states for the study are: Niger, Benue and Plateau. The states that constitutes the north central geopolitical zone are: Kwara, Kogi, Plateau, Nassarawa, Benue and Niger states[8]

2.0 Materials and Methods

2.1 Hidden Markov Model

A Hidden Markov Model (HMM) is a double stochastic process in which one of the stochastic processes is an underlying

Corresponding author: Lawal. A, E-mail: lawal.adamu@futminna.edu.ng, Tel.: +2348036011769

Markov chain which is called the hidden part of the model, the other stochastic process is an observable one. Also a hidden Markov model (HMM) can be considered as a stochastic process whose evolution is governed by an underlying discrete (Markov chain) with a number of states $s_i \in S, i=1, N$, which are hidden, i.e. not directly observable [9]. HMM is characterized by the following:

N = number of states in the model
 M = number of distinct observation symbols per state

Q = state sequence

$$Q = q_1, q_2, q_3, \dots, q_T$$

O = observation sequence

$$O = o_1, o_2, o_3, \dots, o_T$$

Transition probability matrix

$$A = \{a_{ij}\}$$

Observation probability matrix

$$B = \{b_j(o_i)\}$$

Where $b_j(o_i) = p(o_i | q_i = s_j)$

If the observation is continuous a probability density function is used

$$\int_{-\infty}^{+\infty} b_j(x) dx = 1$$

$\pi = \{\pi_j\}$ Initial state probabilities

$\lambda = (A, B, \pi)$ The overall HMM

(1)

2.2 Formulation of the Model

In this research work, our Hidden Markov Model is based on observed data of the rainfall in our study areas. Annual rainfall is considered as state of the HMM while onset, recession, distribution (spread) of the rainfall within a year considered as emission of the HMM. As a result, we make the following assumptions

The transition of annual rainfall state to another state in a year follow a Markov chain of first order dependence as per in [10, 11]

The probability of generating current observation symbol depends only on current state. The

$$P(O | Q, \lambda) = \prod_{t=1}^T P(o_t | q_t, \lambda)$$

(2)

Rainfall is said to start early if it starts within (January-March)

Rainfall is said to start late if it starts on the month of April or beyond

Rainfall is said to stop early if the rainfall does not exceed the month of October of a year

Rainfall is said to stop late if the rainfall exceed the month of October

Let the annual rainfall be modelled by a three state hidden Markov model and eight observations, The states are given below

State1: Low Rainfall (rainfall amount ≤ 800 mm)

State2: Moderate Rainfall ($801 \text{mm} \leq \text{rainfall amount} \leq 1300$ mm)

State3: High Rainfall (Rainfall amount > 1300 mm)

And all the possible observations within a year for all the states are given below:

$A = O_1$ = (rainfall starts early and ends early in that year and it is well spread)

$B = O_2$ = (rainfall starts early and ends early in that year and it is not well spread)

$C = O_3$ = (rainfall starts early and ends late for that year and it is well spread)

$D = O_4$ = (rainfall starts early and ends late for that year and it is not well spread)

$E = O_5$ = (rainfall starts late and ends late for that year and it is well spread)

$F = O_6$ = (rainfall starts late and ends late of that year and it is not well spread)

$G = O_7$ = (rainfall starts late and ends early of that year and it is well spread)

$H = O_8 =$ (rainfall starts late and ends early for that year and it is not well spread)
 The possible transitions between the states and their emissions are illustrated in Figure 1

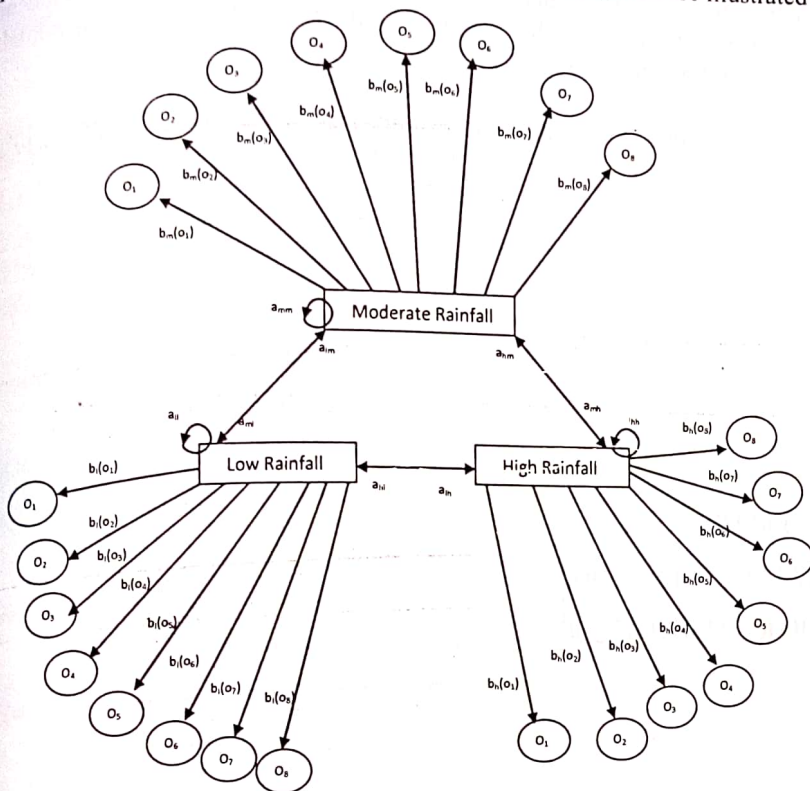


Figure 1: Transition Diagram of the Annual Rainfall Model
 Transition Probability Matrix for The states

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{3}$$

Observation Probability Matrix

$$B = \begin{bmatrix} b_l(o_1) & b_l(o_2) & b_l(o_3) & b_l(o_4) & b_l(o_5) & b_l(o_6) & b_l(o_7) & b_l(o_8) \\ b_m(o_1) & b_m(o_2) & b_m(o_3) & b_m(o_4) & b_m(o_5) & b_m(o_6) & b_m(o_7) & b_m(o_8) \\ b_h(o_1) & b_h(o_2) & b_h(o_3) & b_h(o_4) & b_h(o_5) & b_h(o_6) & b_h(o_7) & b_h(o_8) \end{bmatrix} \tag{4}$$

Initial Probability Distribution

$$\pi = [\pi_1 \quad \pi_2 \quad \pi_3] \tag{5}$$

2.3 Making Predictions With the Model

Likelihood based Prediction is adopted in this work. In this method, the parameters of the model are initialized then trained using Baum-Welch algorithm to attends Maximum likelihood. The forward probability of the training observation sequence is calculated from time $t=1$ to time T using Forward Algorithm [2, 12]. To predict the next state at time $T+1$ and its observation given the present state at time T , we calculate forward probability for each possible observations of the states, then the sequence with highest value of the forward probability at time $T+1$ is taken as predicted state and its observations. The prediction is made for the next two years (at time $T+2$, and at time $T+3$).

3.0 Results and Discussion

3.1 Application of the model in Jos, Plateau State

The data used in this research work, were collected from the archive of the department of Geography and Planning, Faculty of Environmental Sciences, University of Jos, Plateau state, Nigeria. Jos is the capital city of Plateau state from 1977-2015

3.1.1 Validity Test for the Model

Before the model could be used to make prediction the strength of the model need to be tested , to achieve this, we divid dataset into two sets, one training set and the other test set. We estimate the parameter of the test HMM1 using the ra data from 1977-1994, then use it to predict annual rainfall for 1995, 1996 and 1997.

From our dataset we obtained the Transition probability matrix, Observation probability matrix and the Initial probab distribution, as represented by equation (6), (7) and (8)

Transition Probability Matrix

$$A = \begin{bmatrix} 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.5000 & 0.5000 \\ 0.0000 & 0.5714 & 0.4286 \end{bmatrix} \tag{6}$$

Observation probability Matrix

$$B = \begin{bmatrix} 0.0000 & 0.2000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 1.0000 & 0.5000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.5000 \\ 0.0000 & 0.3000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 1.0000 & 0.5000 \end{bmatrix} \tag{7}$$

Initial State probabilities

$$\pi = [0.0556 \quad 0.5000 \quad 0.4444] \tag{8}$$

$$\lambda_1 = (A, B, \pi) \tag{9}$$

Equation (9) is the test hidden Markov model (HMM1)

After 500 iterations of the Baum Welch Algorithm using Matlab, equation (9) stabilised to equation (10)

$$\lambda_1^* = (\hat{A}, \hat{B}, \hat{\pi}) \tag{10}$$

where

$$\hat{A} = \begin{bmatrix} 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.5000 & 0.5000 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \tag{11}$$

$$\hat{B} = \begin{bmatrix} 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.9999 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0001 \\ 0.0000 & 0.6000 & 0.0000 & 0.0667 & 0.0000 & 0.0000 & 0.0667 & 0.2667 \end{bmatrix} \tag{12}$$

$$\hat{\pi} = [0 \quad 0 \quad 1] \tag{13}$$

Comparison of the Predicted States and Observations Using HMM1, and the Actual States and Observations From Dataset

Predicted states and Observations Using the Test model (HMM1)

States 3(1994) → 3(1995) → 3(1996) → 3(1997)
 ↓ ↓ ↓ ↓
 Observations: H B B B

Actual states and Observations from the Dataset

States: 3(1994) → 3(1995) → 3(1996) → 2(1997)
 ↓ ↓ ↓ ↓
 Observations: H H B D

As it can be seen from the prediction time series, the HMM1 was in state 3 (high rainfall) at time T (1994) emitted observation H (rainfall starts late and ends early of that year and it is well spread) then make transition to state 3 (high rainfall) at time T+1 (1995) governed by first order Markov dependence, emitting observation B (rainfall starts early and ends early that year and it is not well spread). Similar interpretation is given to transition to state 3 (high rainfall) at T+2 (1996) and transition to state 3 (high rainfall) at time T+3 (1997) both emitting observation B (rainfall starts early and ends early in that year and it is well spread). The comparison of the predicted states and observations with the actual states and observations from the dataset shows 75% and 50% in states and observation prediction respectively.

3.1.2 Hidden Markov Model for Future Predictions (HMM2)

In order to make predictions for the future years, the whole dataset (rainfall data from 1997 to 2015) was used to estimate the parameters of the model, then make predictions for 2016, 2017 and 2018. From our dataset we obtained equations (14), (15) and (16) respectively.

Transition Probability Matrix

$$A = \begin{bmatrix} 0.0000 & 0.8000 & 0.200 \\ 0.2100 & 0.4211 & 0.3684 \\ 0.0000 & 0.5000 & 0.5000 \end{bmatrix} \tag{14}$$

Observation Probability Matrix

$$B = \begin{bmatrix} 0.0000 & 0.1111 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.2500 \\ 1.0000 & 0.4444 & 0.0000 & 0.5000 & 0.0000 & 0.0000 & 0.0000 & 0.5000 \\ 0.0000 & 0.4444 & 0.0000 & 0.5000 & 0.0000 & 0.0000 & 1.0000 & 0.2500 \end{bmatrix} \tag{15}$$

Initial State Probabilities

$$\pi = [0.1282 \quad 0.4872 \quad 0.3846] \tag{16}$$

$$\lambda_2 = (A, B, \pi) \tag{17}$$

After 500 iterations of the Baum Welch Algorithm, equation (17) stabilised to equation (18)

$$\lambda_2^* = (\hat{A}, \hat{B}, \hat{\pi}) \tag{18}$$

Where

$$\hat{A} = \begin{bmatrix} 0.0000 & 0.6332 & 0.3668 \\ 0.7668 & 0.2332 & 0.0000 \\ 0.0000 & 0.3643 & 0.6357 \end{bmatrix} \tag{19}$$

$$\hat{B} = \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \\ 0.1338 & 0.4137 & 0.0000 & 0.1338 & 0.0000 & 0.0000 & 0.0000 & 0.3187 \\ 0.0000 & 0.9220 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0780 & 0.0000 \end{bmatrix} \tag{20}$$

$$\hat{\pi} = [0 \quad 0 \quad 1] \tag{21}$$

Predicted states and Observations Using HMM2
 States: 3(2015) → 3(2016) → 2(2017) → 1(2018)

Observations: B B B B
 The prediction time series shows that, the annual rainfall is in state 3 (high rainfall) at time T (2015) with observation B (rainfall starts late and ends early of that year and it is not well spread) will make transition to state 3 (high rainfall) at time T+1 (2016) according to first order Markov dependence, emitting observation B (rainfall starts early and ends early in that year and it is not well spread), it then make transition to state 2 (moderate rainfall) at time T+2 (2017) and later to state 1 (low rainfall) at time T+3 (2018) both emitting Observation B (rainfall starts early and ends early in that year and it is not well spread).

3.2 Application of the model in Makurdi, Benue State

The Rainfall data used in this model was collected from the archive of Nigerian Meteorological agency, Maitama, A from 2005 -2015.

From the dataset we obtained equation (22) , (23) and (24)

Transition Probability Matrix

$$A = \begin{bmatrix} 0.3333 & 0.3333 & 0.3333 \\ 0.1111 & 0.4444 & 0.4444 \\ 0.1428 & 0.5714 & 0.2857 \end{bmatrix} \tag{22}$$

Observation Probability Matrix

$$B = \begin{bmatrix} 0.3333 & 0.1666 & 0.3333 & 0.0909 & 0.3333 & 0.3333 & 0.3333 & 0.3333 \\ 0.3333 & 0.5000 & 0.3333 & 0.5454 & 0.3333 & 0.3333 & 0.3333 & 0.3333 \\ 0.3333 & 0.3333 & 0.3333 & 0.3636 & 0.3333 & 0.3333 & 0.3333 & 0.3333 \end{bmatrix} \tag{23}$$

Initial State Probabilities

$$\pi = [0.0000 \quad 0.6360 \quad 0.3636] \tag{24}$$

$$\lambda_3 = (A, B, \pi) \tag{25}$$

After 400 iterations of the Baum Welch Algorithm, equation (25) stabilised to equation (26)

$$\lambda_3^* = (\hat{A}, \hat{B}, \hat{\pi}) \tag{26}$$

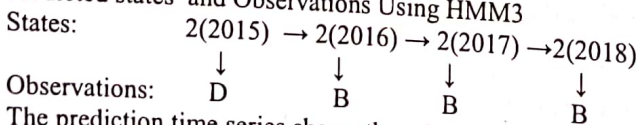
Where ,

$$\hat{A} = \begin{bmatrix} 0.0000 & 0.0000 & 1.0000 \\ 0.0000 & 0.5000 & 0.5000 \\ 1.0000 & 0.0000 & 0.0000 \end{bmatrix} \tag{27}$$

$$\hat{B} = \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.2000 & 0.8000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix} \tag{28}$$

$$\hat{\pi} = [0 \quad 1 \quad 0] \tag{29}$$

Predicted states and Observations Using HMM3



The prediction time series shows that, the annual rainfall is in state 2 (moderate rainfall) at time T(2015) with observation (rainfall starts early and ends late of that year and it is not well spread) will make transition to state 2 (moderate rainfall) time T+1(2016) according to first order Markov dependence, emitting observation B(rainfall starts early and ends early that year and it is not well spread), it then make transition to state2 (moderate rainfall) at time T+2 (2017) and later to state (moderate rainfall) at time T+3(2018) both emitting Observation B (rainfall starts early and ends early in that year and it not well spread).

3.3 Application of the model in Maikunkele, Niger State

The Rainfall data used in this model was collected from the archive of Nigerian Meteorological agency, Minna, international Airport, Maikunkele, Niger State from 1980-2015. In order to make predictions for the future years, the whole dataset (rainfall data from 1980 to 2015) was used to estimate the parameters of the model, then make predictions for 2016, 2017 and 2018. From the dataset we obtain equation (30), (31) and (33)

Transition Probability Matrix

$$A = \begin{bmatrix} 0.3333 & 0.3333 & 0.3333 \\ 0.0340 & 0.7240 & 0.2413 \\ 0.0833 & 0.5833 & 0.3333 \end{bmatrix}$$

(30)

Observation Probability Matrix

$$B = \begin{bmatrix} 0.2500 & 0.0714 & 0.3333 & 0.0909 & 0.2500 & 0.2500 & 0.1428 & 0.0769 \\ 0.2500 & 0.7142 & 0.3333 & 0.8180 & 0.5000 & 0.5000 & 0.4286 & 0.6923 \\ 0.5000 & 0.2140 & 0.3333 & 0.0909 & 0.2500 & 0.2500 & 0.4286 & 0.2307 \end{bmatrix}$$

(31)

Initial State Probabilities

$$\pi = [0.0000 \quad 0.8055 \quad 0.1944]$$

(32)

$$\lambda_4 = (A, B, \pi)$$

(33)

After 300 iterations of the Baum Welch Algorithm, equation (33) stabilised to equation (34)

$$\lambda_4^* = (\hat{A}, \hat{B}, \hat{\pi})$$

(34)

Where,

$$\hat{A} = \begin{bmatrix} 0.0000 & 0.0000 & 1.0000 \\ 0.0000 & 0.8621 & 0.1379 \\ 0.2000 & 0.8000 & 0.0000 \end{bmatrix}$$

(35)

$$\hat{B} = \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.3667 & 0.0000 & 0.2667 & 0.0000 & 0.0333 & 0.0000 & 0.3333 \\ 0.2000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.8000 & 0.0000 \end{bmatrix}$$

(36)

$$\hat{\pi} = [0 \quad 1 \quad 0]$$

(37)

Predicted states and Observations Using HMM4

States: 2(2015) → 2(2016) → 2(2017) → 2(2018)

Observations: D B B B

The prediction time series shows that, the annual rainfall is in state 2(moderate rainfall) at time T(2015) with observation D (rainfall starts early and ends late of that year and it is not well spread) will make transition to state 2 (moderate rainfall) at time T+1 (2016) according to first order Markov dependence, emitting observation B(rainfall starts early and ends early in that year and it is not well spread), it then make transition to state2 (moderate rainfall) at time T+2 (2017) and later to state2(moderate rainfall) at time T+3 (2018) both emitting Observation B (rainfall starts early and ends early in that year and it is not well spread)

Table 1: Comparison of the Result from the Three Locations

		States and Observations From the Dataset										Predicted states and observations			
		2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Plateau	States	2	2	1	2	2	1	2	2	3	3	3	3	2	1
	Observations	H	B	H	H	H	H	H	H	H	B	B	B	B	B
Niger	States	2	3	3	2	3	2	2	3	2	2	2	2	2	2
	Observations	H	B	B	H	H	F	D	H	B	G	D	B	B	B
Benue	States	2	3	3	2	3	2	2	3	2	2	2	2	2	2
	Observations	B	B	D	D	D	D	B	D	D	D	D	B	B	B

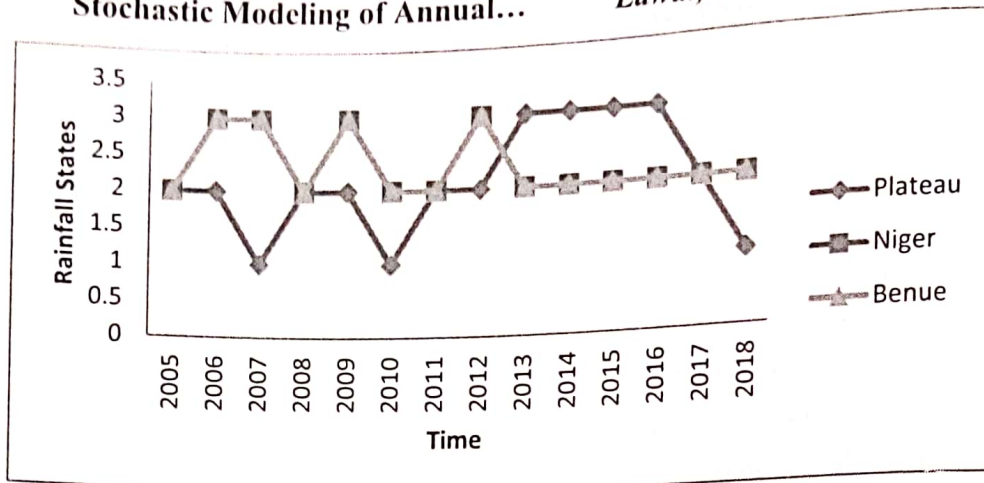


Figure 2: Graph of the results for the three locations

From Table 1 and Figure 2, it can be observed that Benue state and Niger state almost have the same pattern of rainfall year. Both states, experienced more of state 2 (moderate rainfall). The difference between the pattern of rainfall of these locations is in the observations. Benue state experienced more of observation D than any other observation. Interpretation of this, is that, rainfall tends to start earlier and ends late in year. However Niger state experiences more observation B, which means that, rainfall start earlier and always ends earlier in the year. In general, Benue state experiences longer period of rainfall in the years as compared to Niger state, but the two states almost have the same amount of rainfall the years. Plateau state has different pattern of annual rainfall as compared to the two other states. All classes of the rainfall are experienced in plateau state but high rainfall is most common. It experienced more of the observation H, which means that, rainfall tends to start late and end early in a plateau state. We may wish to suggest that, the almost the same pattern of rainfall of Niger and Benue states may be one of the reasons why both states are leading producers of food in the country especially tuber crop like Yam.

4.0 Conclusion

A hidden Markov model to predict rainfall onset, recession, amount and distribution (spread) has been presented in this paper. The validity test showed 75% and 50% accuracy in states and observations prediction respectively. The model was implemented in three selected states of the north central Nigeria (Plateau, Benue and Niger state). The result shows that Benue and Niger states have almost the same pattern of rainfall. Both states experienced more of moderate rainfall, but the difference between the two locations is in the observations. Plateau has different pattern of rainfall. It experiences all classes of the rainfall but experienced more of high rainfall and observation H (rainfall starts late and ends early for that year) and it is not well spread. The study therefore suggests that, the similarity of the pattern of rainfall of Niger state and Benue state may be one of the reasons responsible that the two states are leading producers of food crops especially tuber crop like yam in Nigeria.

5.0 References

- [1] Iloje N.P. (1981). A New Geography of Nigeria, Longman Nigeria Limited.
- [2] Rabiner, L.R.(1989). A Tutorial on hidden Markov models and Selected Applications in Speech Recognition. Proc. IEEE, 77, . 257-286.
- [3] Durbin. R, S. Eddy, A. Krogh, and G. Mitchison(1998) Biological sequence analysis. Probabilistic models for proteins and nucleic acids. Cambridge University Press, 1998
- [4] Vaseghi, S.V.(2006).Advanced Digital Signal Processing and NoiseReduction.4th Edition, Wiley
- [5] Baldi, P., & Brunak, S. (2001). Bio informatics. A Bradford Book, The MIT Press, Cambridge.
- [6] Hirsch. H-G(2001). HMM adaptation for applications in telecommunication. Speech Communication, Vol.34,iss 1-2, 127-139
- [7] Fink, G.A. (1989). MarkovModels for PatternRecognition.From Theory toApplications, Springer
- [8] Blessing Smart (2015). Top 6 Geopolitical zones in Nigeria. Information guide in Nigeria. <http://infoguidenigeria.com/top-6-geopolitical-zones-nigeria>. Accessed 27/12/2016.
- [9] Enza, Messina, & Daniele, Toscani (2007). Hidden Markov models for scenario generation. Journal of Management Mathematics, 19, 379-40
- [10] Lawal, Adamu (2013). The Use of Markov Model in Predicting the Rainfall Pattern for Crop Production. Unpublished M.Tech Thesis, Department of Mathematics and Statistics, Federal University of Technology, Minna, Nigeria
- [11] Ross S.M.(1989). Introduction to Probability Models Academic Press, Ltd, London.
- [12] Hyang-Won Lee (2014). Hidden Markov Models, Probability and Statistics. Spring, 1-18