Knowing the nature of rainfall distribution before each growing season begins in most parts of the world, especially in developing countries has always been a fundamental problem to the farmers, and this has over the years led to improper crop planning and cultivation, consequently led to poor harvest. This book demonstrates the application of Hidden Markov Model (HMM) in rainfall pattern prediction for the purpose of crop production using empirical data. The validity tests for the models showed that they are reliable and dependable. Therefore, results from these models could serve as a guide to the farmers and the government to plan strategies for high crop production in the region. The results from the models could also assist the residents to better understand the dynamics of rainfall which may be helpful for effective planning and viable crop production.

Dr. Lawal Adamu is from Fiche-Kuchi, Paiko, Niger State, Nigeria. He has First Degree in Mathematics and Computer Science, Masters Degree in Mathematics and PhD Degree in Applied Mathematics. He is a Lecturer with Department of Mathematics Federal University of Technology, Minna, Nigeria. His Research Area is Optimization Theory.

978-613-4-98532-1

Lawal Adamu

Hidden Markov Model and Its Application in Rainfall Pattern Prediction

LAP LAMBERT Academic Publishing

Lawal Adamu

Hidden Markov Model and Its Application in Rainfall Pattern Prediction

Lawal Adamu

Hidden Markov Model and Its Application in Rainfall Pattern Prediction

LAP LAMBERT Academic Publishing

Imprint
Any brand names and product names mentioned in this book are subject to trademark, brand or patent protection and are trademarks or registered trademarks of their respective holders. The use of brand names, product names, common names, trade names, product descriptions etc. even without a particular marking in this work is in no way to be construed to mean that such names may be regarded as unrestricted in respect of trademark and brand protection legislation and could thus be used by anyone.

Cover image: www.ingimage.com

Publisher:
LAP LAMBERT Academic Publishing
is a trademark of
International Book Market Service Ltd., member of OmniScriptum Publishing Group
17 Meldrum Street, Beau Bassin 71504, Mauritius

Printed at: see last page
ISBN: 978-613-4-98532-1

MONOGRAPH

HIDDEN MARKOV MODEL AND ITS APPLICATION IN RAINFALL PATTERN PREDICTION

LAWAL ADAMU(PhD)

DEPARTMENT OF MATHEMATICS

FEDERAL UNIVERSITY OF TECHNOLOGY MINNA

NIGER STATE, NIGERIA.

PREFEACE

Hidden Markov Model (HMM) is one of the best and effective pattern recognition stochastic mathematical model currently in existence. What make it effective and unique is its double stochastic nature. Its effectiveness and benefits is not known to many researchers, this book is an attempt to provide such information to the intending researchers.

This book is made up of two chapters: Chapter 1 discusses the basic concept of HMM while Chapter 2 presents applications of HMM in rainfall pattern prediction for the purpose of crop production in some selected states of North Central Nigeria.

ACKNOWLEDGEMENTS

Praise be to Allah the Lord of all worlds, the Most gracious, the Most merciful, the Most knowledgeable and the Wise, for granting me the opportunity to write this monograph.

I wish to acknowledge my mentor Prof. U. Y Abubakar, who introduced me to Stochastic Mathematical modelling.

My appreciation also goes to Dr. D. Hakimi for his support and constructive advice which have added value to this work.

I wish to also appreciate the efforts and support of Prof. Andrew Saba Gana of the Department of Crop Production, Federal University of Technology, Minna.

I also want to acknowledge Mallama Maimuna of Nigerian Meteorological Agency, and the other staff of the Agency that provided the rainfall data used in this research work. The sensitive and critical advice of: Mallam Lawal Mohammed the Deputy Director Agricultural Services, Mallam Yakubu. A . Ibrahim the Senior Planning Officer, other staff and Management of the Niger State Agricultural Mechanization and Development Authority are well appreciated and acknowledged.

I would also want to acknowledge the support of: Mallam Aliyu Shehu, Audu Adamu, Hauwa Adamu and Mallam Hashimu Musa, for their support and encouragement during the course of this work

My profound gratitude goes to my Late Uncle Mallam Adamu Machi, for giving me the best future.

My ultimate thanks go to my wife Zaliya Lawal Waziri and to my children: Adamu, Deenat and Barakatu for their support and love.

TABLE OF CONTENTS

Content **Page**

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

1.0 Markov Models

Markov models are stochastic models with a Markov property. A model is said to have a Markov property if the future state of the process depends on the current state and not on its past history. Markov model was first introduced by a Russian mathematician named Andrey Andreyevich Markov in 1906 when he produced the first theoretical results for stochastic processes by using the term "chain" for the first time. His research area later became known as Markov process and Markov chains. A generalization to countable infinite state space was given by kolmogorov.

1.1 Markov Chain

Consider a stochastic process $\{X_n, \quad n = 0,1,2,3,........\}$ which takes on a finite or countable number of possible values. If $X_n = i$ then the process is in state *i* at time ***n*.** Whenever the process is in state *i,* there is a fixed probability P_{ij} that it will next be in state j. Thus:

$$P\{X_{n+1} = j \mid X_0 = i_0, ..., X_{n-2} = i_{n-2}, X_{n-1} = i_{n-1}, X_n = i\} = P_{ij} \tag{1.1}$$

for all states $i_0, i_i, ..., i_{n-1}, i, j \quad and\ all\ n \geq 0$. This type of a stochastic process is called Markov chain**.** A Markov chain has a "memoryless" property that is, the future state of the process depends on the present and not on the past.

1.1.2 Chapman-Kolmogorov Equations (CKE)

We 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. That is,

$$P_{ij}^n = P\{X_{n+m} = j \mid X_m = i\}, \quad n \geq 0, \ i, j \geq 0 \tag{1.2}$$

Of course $P_{ij}^1 = P_{ij}$. The CKE provide a way for computing these n-step transition probabilities. Thus

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m \quad for\ all\ n, m \geq 0,\ all\ i, j \tag{1.3}$$

$P_{ik}^{n}P_{kj}^{m}$ represents the probability that starting in state i the process will go to state j in $n+m$ transitions through a path which takes it into state k at the nth transition. Hence, summing over all intermediate states k yields the probability that the process will be in state j after n + m transitions. We have

$$\left.\begin{aligned}
P_{ij}^{n+m} &= P\{X_{n+m}=j \mid X_0=i\} \\
P_{ij}^{n+m} &= \sum_{k=0}^{\infty} P\{X_{n+m}=j, X_n=k \mid X_0=i\} \\
P_{ij}^{n+m} &= \sum_{k=0}^{\infty} P\{X_{n+m}=j \mid X_n=k, X_0=i\}P\{X_n=k \mid X_0=i\} \\
P_{ij}^{n+m} &= \sum_{k=0}^{\infty} P_{kj}^{m}P_{ik}^{n}
\end{aligned}\right\} \qquad (1.4)$$

Now, let $P^{(n)}$ represents the matrix of n-step transition probabilities P_{ij}^{n}, then Equation (1.3) asserts that $P^{(n+m)} = P^{(n)} \cdot P^{(m)}$.The dot represents matrix multiplication. Hence, in particular, $P^{(2)} = P^{(1+1)} = P \cdot P = P^{2}$ Equation (1.5) is obtained by induction,

$$P^{(n)} = P^{(n-1+1)} = P^{n-1} \cdot P = P^{n} \qquad (1.5)$$

This means that, the n-step transition matrix is obtained by multiplying the matrix P by itself n times.

1.1.3 Limiting Probabilities

If $n \to \infty$ in equation (1.5), the probability will converge to some value, which is the same for all i. That is, there exist 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

1.2 Hidden Markov Model

A Hidden Markov Model (HMM) is a double stochastic process in which one of the stochastic processes is an underlying Markov chain which is called the hidden part of the model, the other stochastic process is an observable one. Also a HMM can be considered as a stochastic process whose evolution is governed by an underlying discrete (Markov chain) with a finite number of states which are hidden, i.e. not directly observable (Enza and Daniele, 2007). A HMM consists of two stochastic processes. The first stochastic process is a Markov chain that is characterized by states and transition probabilities. The states of the chain are externally not visibly therefore " hidden". The second stochastic process produces emissions observable at each moment, depending on a state-dependent probability distribution. Hidden Markov model is very influential in stochastic world because of its uniqueness and applicable mathematical structure and its independence assumption between the consecutive observations, motivates further applications. Hidden Markov Models have been successfully applied in automatic speech recognition and speech synthesis (Rabiner, 1989), molecular biology for DNA and protein sequencing (Durbin *et al*., 1998), signal processing (Vaseghi, 2006), bioinformatics (Baldi *et al*., 2001), telecommunication (Hirsch, 2001) and in pattern recognition (Fink, 1989).

1.2.1 Characteristics of Hidden Markov Model

Hidden Markov Model is characterized by the following

N = number of states in the model

M = number of distinct observation symbols per state

Q = state sequence

$$Q = q_{1,} q_2, q_3, \ldots\ldots\ldots\ldots q_T \tag{1.6}$$

O = observation sequence

$$O = o_{1,} o_{2,} o_3 \ldots\ldots\ldots\ldots o_T \tag{1.7}$$

Transition probability matrix $A = \{a_{ij}\}$ (1.8)

Observation probability matrix $B = \{b_j(o_t)\}$ (1.9)

where $b_j(o_t) = p(o_t \mid q_t = s_j)$

If the observation is continuous a probability density function is used as follows:

$$\int_{-\infty}^{+\infty} b_j(x)dx = 1 \tag{1.10}$$

Initial state probabilities $\pi = \{\pi_j\}$ (1.11)

The overall HMM $\lambda = (A, B, \pi)$ (1.12)

1.2.2 Problems of Hidden Markov Model

According to Rabiner (1989), the Problems of HMM are as follows:

Problem 1

Given the HMM $\lambda = (A, B, \pi)$, what is the probability of generating a specific observation sequence

$$O = \{o_1, o_{2,...,} o_T\}. \tag{1.13}$$

Now the problem of computing $P(O \mid \lambda)$ arises

Problem 2

Given the Observation Sequence $O = \{o_1, o_{2,...,} o_T\}$ and the model $\lambda = (A, B, \pi)$

How do we determine the optimal states sequence?

$$Q = \{q_1, q_2, ..., q_T\} \tag{1.14}$$

Problem 3

Given the Observation sequence $O = \{o_1, ..., o_T\}$, How to estimate the parameters $\lambda = (A, B, \pi)$ of the HMM?. This is to find the model $\lambda^* = (A, B, \pi)$ that maximizes the probability of O.

1.2.3 Solutions to Hidden Markov Model Problems

1.2.3.1 Solution to the problem1

Let $\lambda = (A, B, \pi)$ be a given HMM model and let $O = \{o_1, o_2, ..., o_T\}$ be the observation sequence and $Q = \{q_1, q_2, ..., q_T\}$ be the state sequence, we are to find $P(O \mid \lambda)$.Then by definition of B we have that

$$P(O \mid Q, \lambda) = P(o_T \mid o_1, ..., o_{T-1}, Q, \lambda) P(o_1, ..., o_{T-1} \mid Q, \lambda) \tag{1.15}$$

$$P(O|Q,\lambda) = P(o_T \mid q_T,\lambda)P(o_1,...,o_{T-1} \mid Q,\lambda)$$

.

.

.

$$\prod_{t=1}^{T} P(o_t \mid q_t,\lambda) = b_{q_1}(o_1)b_{q_2}(o_2)....b_{q_T}(o_T) \qquad (1.16)$$

And by the definition of Π *and* A it follows that

$$P(Q|\lambda) = P(q_1)\prod_{t=2}^{T} P(q_t \mid q_{t-1}) = \pi_{q_1} a_{q_1 q_2} a_{q_2 q_3} ...a_{q_{T-1} q_T} \qquad (1.17)$$

Using the principle of joint probability, we obtain

$$P(O|Q,\lambda)P(Q|\lambda) = \frac{P(O \cap Q \cap \lambda)}{P(Q \cap \lambda)} . \frac{P(Q \cap \lambda)}{P(\lambda)} = \frac{P(Q \cap Q \cap \lambda)}{P(\lambda)} \qquad (1.18)$$

$$P(O,Q|\lambda) = P(O|Q,\lambda)P(Q|\lambda) \qquad (1.19)$$

$$\begin{aligned} P(O,Q|\lambda) &= P(q_1)\prod_{t=2}^{T} P(q_t \mid q_{t-1})\prod_{t=1}^{T} P(o_t \mid q_t) \\ &= \pi_{q_1} b_{q_1}(o_1) a_{q_1 q_2} b_{q_2}(o_2).....a_{q_{T-1} q_T} b_{q_T}(o_T) \end{aligned} \qquad (1.20)$$

Equation (1.20) is the joint probability of equation (1.16) and equation (1.17)

Thus we have

$$P(O|\lambda) = \sum_{all\ possipleQ} P(O,Q|\lambda) = \sum_{Q} P(O|Q,\lambda)P(Q|\lambda) \qquad (1.21)$$

This direct computation is generally infeasible, since it requires $2TN^T$ Multiplication and as result of this limitation, we employ the use of Forward Algorithm. The Forward Algorithm involves the use of recursive formula to compute $P(O|\lambda)$ efficiently.

Now, let $\alpha_t(i)$ be the probability of observing the partial sequence $\{o_1, o, ..., o_t\}$ until time t and being in state s_i at time t, given the model λ, That is,

$$\alpha_t(i) = P(o_1, o_2, ..., o_t, q_t = s_i \mid \lambda) \tag{1.22}$$

The probability can be calculated recursively as:

Initialization: $\alpha_1(i) = P(o_1, q_1 = s_i \mid \lambda)$

$$\alpha_1(i) = \pi_i b_i(o_1) \tag{1.23}$$

The Recursion

$$\alpha_{t+1}(j) = P(o_1, o_2, ..., o_{t+1}, q_{t+1} = s_j \mid \lambda) \tag{1.24}$$

$$= P(o_1, o_2, ..., o_{t+1} \mid q_{t+1} = s_j, \lambda) P(q_{t+1} = s_j \mid \lambda)$$

$$= P(o_1, o_2, ..., o_t \mid q_{t+1} = s_j, \lambda) P(o_{t+1} \mid q_{t+1} = s_j, \lambda) P(q_{t+1} = s_j \mid \lambda)$$

$$= P(o_1, o_2, ..., o_t, q_{t+1} = s_j \mid \lambda) b_j(o_{t+1})$$

$$= b_j(o_{t+1}) \sum_i P(o_1, o_2, ..., o_t, q_t = s_i, q_{t+1} = s_j \mid \lambda)$$

$$= b_j(o_{t+1}) \sum_i P(o_1, o_2, ..., o_t, q_{t+1} = s_j \mid q_t = s_i, \lambda) P(q_t = s_i \mid \lambda)$$

$$= b_j(o_{t+1}) \sum_i P(o_1, o_2, ..., o_t \mid q_t = s_i, \lambda) P(q_{t+1} = s_j \mid q_t = s_i, \lambda) P(q_t = s_i \mid \lambda)$$

$$= b_j(o_{t+1}) \sum_i P(o_1, o_2, ..., o_t \mid q_t = s_i, \lambda) q_{ij} P(q_t = s_i \mid \lambda)$$

$$= b_j(o_{t+1}) \sum_i P(o_1, o_2, ..., o_t, q_t = s_i \mid \lambda) q_{ij}$$

$$\alpha_{t+1}(j) = \sum_i \alpha_t(i) q_{ij} b_j(o_{t+1}) \tag{1.25}$$

From equation (1.25) we can see that the probability of the observation sequence can be calculated as:

$$P(O|\lambda)=\sum_{i=1}^{N}P(O,q_T=s_i|\lambda) \tag{1.26}$$

$$P(O|\lambda)=\sum_{i=1}^{N}\alpha_T(i) \tag{1.27}$$

Similarly, we define the backward variable, $\beta_t(i)$ as the probability of being in state s_i at time t and observing the partial sequence, that is, $o_{t+1},...,o_T$

$$\beta_t(i)=P(o_{t+1},...,o_T|q_t=s_i,\lambda) \tag{1.28}$$

Initialization: $\beta_T(i)=1$

Recursively, we obtain

$$\beta_t(i)=P(o_{t+1},...,o_T|q_t=s_i,\lambda)$$

$$\beta_t(i)=\sum_j P(o_{t+1},...,o_T,q_{t+1}=s_j|q_t=s_i,\lambda) \tag{1.29}$$

$$\beta_t(i)=\sum_j P(o_{t+1},...,o_T|q_{t+1}=s_j,\lambda)P(q_{t+1}=s_j|q_t=s_i,\lambda)$$

$$\beta_t(i)=\sum_j P(o_{t+1}|q_{t+1}=s_j,q_t=s_i,\lambda)P(o_{t+2},...,o_T|q_{t+1}=s_j,q_t=s_i,\lambda)a_{ij}$$

$$\beta_t(i)=\sum_j P(o_{t+1}|q_{t+1}=s_j,\lambda)P(o_{t+2},...,o_T|q_{t+1}=s_j,\lambda)a_{ij}$$

$$\beta_t(i)=\sum_j P(o_{t+1}|q_{t+1}=s_j,q_t=s_i,\lambda)P(o_{t+2},...,o_T|q_{t+1}=s_j,q_t=s_i,\lambda)a_{ij}$$

$$\beta_t(i)=\sum_j a_{ij}b_j(o_{t+1})\beta_{t+1}(j) \tag{1.30}$$

1.2.3.2 Solution to the problem 2 (Decoding)

The goal of the decoding is to find the states sequence

$Q = \{q_1, q_2, ..., q_T\}$ having the highest probability of generating the observation sequence $O = \{o_1, o_{2,} ..., o_T\}$, given the model λ

Let $\gamma_t(i)$ be the probability of being in state s_i at time t, given O and λ. This can be computed as

$$\gamma_t(i) = P(q_t = s_i \mid O, \lambda) \tag{1.31}$$

$$= \frac{P(O \mid q_t = s_i, \lambda) P(q_t = s_i \mid \lambda)}{P(O \mid \lambda)}$$

$$= \frac{P(o_1 ... o_t \mid q_t = s_i, \lambda) P(o_{t+1} .... o_T \mid q_t = s_i, \lambda) P(q_t = s_i, \lambda)}{\sum_j P(O, q_t = s_j \mid \lambda)}$$

$$= \frac{P(o_1 ... o_t, q_t = s_i \mid \lambda) P(o_{t+1} ... o_T \mid q_t = s_i, \lambda)}{\sum_j P(O \mid q_t = s_j, \lambda) P(q_t = s_j \mid \lambda)}$$

$$\therefore \gamma_t(i) = \frac{\alpha_t(i) \beta_t(i)}{\sum_j \alpha_t(j) \beta_t(j)} \tag{1.32}$$

Decoding Using Viterbi Algorithm

To find a best states sequence $Q = \{q_1, q_2, ..., q_T\}$ for a given observation sequence

$O = \{o_1, o_2, ..., o_T\}$. The Viterbi algorithm is used, it has advantage of less computation and it is generally efficient. The Viterbi algorithm can be viewed as a dynamic programming algorithm applied to the HMM or as a modified algorithm. The Viterbi algorithm picks and remembers the best path, instead of summing up probabilities from different paths coming to the same destination state. To achieve this, We define a new variable.

$$\delta_t(i) = Max P[q_1, q_2 ... q_{t-1}, q_t = i, o_1, o_2, ..., o_t \mid \lambda] \tag{1.33}$$

This represents the best score along a single path at time t, which accounts for the first t observation and ends in state i.. This probability can be calculated based on the dynamic programming and the optimal path can be retrieved by backtracking from T.

Initialization:

$$\delta_1(i) = \pi_i b_i(o_1) \quad , \; \psi_1(i) = 0 \quad \forall_i$$

Recursion $(t = 2.................T)$

$$\delta_t(j) = \max_i \delta_{t-1}(i) a_{ij} b_j(o_t) \quad , \quad \psi_t(j) = \arg\max_i \delta_{t-1}(i) a_{ij} \tag{1.34}$$

Termination $q_T^* = \arg\max_i \delta_T(i)$

Path (state sequence) backtracking

$$q_T^* = \psi_{t+1}(q_{t+1}^*), \quad t = T-1, T-2, ..., 1 \tag{1.35}$$

The variable $\psi_t(j)$ keeps track of the optimal state at time t-1 if the state at time t is j. Once the best state at time T is known (which is q_T^*), the optimal path can be retrieved by backtracking the variable ψ

1.2.3.3 Solution to the problem 3

This is a learning process and it involves the adjustment of the model parameters to best fit the observations. In general, the goal of learning is to calculate λ^* that maximizes the likelihood $P(O \mid \lambda)$ of the sample of training sequence, we define $\gamma_t(i, j)$ as the probability of being s_i at time t and s_j at time $t+1$, given the whole observation O and the model λ , that is

$$\gamma_t(i, j) = P(q_t = i, q_{t+1} = s_j \mid O, \lambda) \tag{1.36}$$

$$= \frac{P(O / q_t = s_i, \quad q_{t+1} = s_j, \lambda) P(q_t = s_i, q_{t+1} = s_j \mid \lambda)}{P(O \mid \lambda)}$$

$$= \frac{P(O \mid q_t = s_i, \quad q_{t+1} = s_j, \lambda) P(q_{t+1} = s_j \mid q_t = s_i, \lambda) P(q_t = s_i \mid \lambda)}{P(O \mid \lambda)}$$

$$= \frac{P(o_1, ..., o_t \mid q_t = s_i, \lambda) P(o_{t+1} \mid q_{t+1} = s_j, \lambda) P(o_{t+1}, ..., o_T \mid q_{t+1} = s_j, \lambda) a_{ij} P(q_t = s_j \mid \lambda)}{P(O \mid \lambda)} \quad (1.37)$$

$$= \frac{P(o_1, ..., o_t, q_t = s_i \mid \lambda) P(o_{t+1} \mid q_{t+1} = s_j, \lambda) P(o_{t+1}, ..., o_T \mid q_{t+1} = s_j, \lambda) a_{ij}}{\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)} \quad (1.38)$$

$$\therefore \gamma_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_t(i) a_{ij} b_l(o_{t+1}) \beta_{t+1}(j)} \quad (1.39)$$

Also the probability of being s_i at time t and s_j at time $t+1$, given the whole observation O can be obtained by

$$\gamma_t(i) = \sum_{j=1}^{N} \gamma_t(i, j) \quad (1.40)$$

$\gamma_t(i)$ is the expected number of visits to state s_i at time t, and $\gamma_t(i, j)$ is the expected number of transitions from s_i (at time t) to s_j (at time t+1).

Using equations (1.12), (1.22), (1.28), (1.31) and (1.36), the Baum-Welch Algorithm is as follows:

The Baum-Welch Algorithm

1. Initialize $\lambda = (A, B, \pi)$
2. Calculate $\alpha_t(i) \quad and \quad \beta_t(i) \quad for \quad all \quad t, i$
3. Calculate $\gamma_t(i, j) \quad and \quad \gamma_t(i) \quad for \quad all \quad t$
4. Estimates a λ as follow

$$\hat{\pi}_i = \gamma_1(i) \tag{1.41}$$

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \gamma_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \tag{1.42}$$

$$\hat{b}_j(m) = \frac{\sum_{t=1}^{T} \gamma_t(j)\{o_t = v_m\}}{\sum_{t=1}^{T} \gamma_t(j)} \tag{1.43}$$

Repeat Step 2-4

where

$$\sum_{t=1}^{T-1} \gamma_t(i) = \text{expected number of transitions from state } i$$

$$\sum_{t=1}^{T-1} \gamma_t(i,j) = \text{expected number of transitions from state } i \text{ to state } j$$

With the above definition, one can then outline the Baum-Welch Re-estimation formula as follows

$$\hat{\pi}_i = \text{expected frequency in state } i \text{ at time } t = 1$$
$$= \gamma_1(i) \tag{1.44}$$

$$\hat{a}_{ij} = \frac{\text{expected number of } \textit{transitions} \text{ from state i to state j}}{\text{expected number of } \textit{transitions} \text{ from state i}} \tag{1.45}$$

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \gamma_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}$$

$$\hat{b}_j(m) = \frac{\text{expected number of times in state } j \text{ and observing } V_m}{\text{expected number of times in state } j} \tag{1.46}$$

$$\hat{b}_j(m) = \frac{\sum_{t=1}^{T} \gamma_t(j)\{o_t = v_m\}}{\sum_{t=1}^{T} \gamma_t(j)}$$

CHAPTER TWO

2.0 Application Of The Hidden Markov Model in Rainfall Pattern Prediction For Crop Production

Rainfall Modelling and predictions are essential input in agricultural production and management of water resources. Bumper harvest and lean years depend largely on rainfall variability and quantity. Rainfall exhibits a strong variability in time and space across the globe. It is well established that rainfall is changing on both the global and the regional scale due to global warming (Hulme *et al*., 1998). Rainfall is the principal phenomenon driving many hydrological extremes such as droughts, floods, debris, landslides and mud-flows; its modelling and analysis are typical problems in applied hydrometeorology (Barkotulla, 2010). Hence, its stochastic modelling is imperative for the avoidance of natural disaster. Having a well established knowledge of rainfall pattern of an area is an important information for future planning. This is applicable in areas like: industry, agriculture, insurance, water resources management and the entire planning of a country economy. Subsequently, proper information on rainfall pattern is fundamental for the design of water supply management, complementary irrigation schemes and the evaluation of alternative cropping system for effective soil water management plans (Barkotulla, 2010). Such information could be beneficial in determining the best adapted plant species and the optimum time of seedling to re-establish vegetation on deteriorated rangelands. Rainfall modelling and prediction are becoming increasingly in demand, because of the uncertainty that is involved with rainfall and the dependence of agricultural production on rainfall. Annual rainfall varies in Nigeria. It decreases from the south to the north where rainfall comes late and not evenly spread across the rainy season and consequently has adverse effect on agricultural activities (Iloeje, 1981). The selected locations for the application of the HMM are Niger and Plateau states in north central Nigeria. Rainfall in the north central Nigeria, is not ordinarily predictable, its occurrence and amount varies from year to year. 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. This 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.

2.1 Seasonal Rainfall Modelling for Crop Production in Maikunkele Niger state, Nigeria

Niger State is the biggest state in the middle belt of Nigeria with an area of about 76,363 km^2. It is located in 9.6490° N, 6.4530° E. It has Guinea savannah vegetation which covers the entire landscape and is characterized by tall grasses woodlands and interspersed with tall dense species. It has tropical climate. Niger Sate is an agriculture-based state in Nigeria, with about 80% of its population engaged in agricultural activities (Blessing Smart, 2015). Rainfall is the major source of water for agricultural production in Niger state (Lawal Adamu, 2013). Knowing the nature of rainfall distribution before each growing season begins in this apart of Nigeria has always been a fundamental problem to the farmers, and this has over the years led to improper crop planning and cultivation, consequently led to poor harvest. In this section, HMM will be use to predict seasonal rainfall pattern for crop production. This will provide the farmers with information that will enable them plan strategies for high crop production in region.

2.1.1 Formulation of the Model

In this model, classification of the states (seasons) is be based on our study area, planting/growing and maturity period of the major crops being grown in the study area. The classification of states for a Markov model is usually based on the purpose in which the model is intended to achieve. The purpose of this model is to identify the major crops being grown in the study area and their maturity period in relation to rainfall. This will make it possible to provide a quantitative prediction of rainfall pattern to the farmers and policy makers to boost crop production.

Model Assumptions

i. The transition of rainfall state to another state follows a Markov chain of the first order dependence, as represented by equation (1.1)

ii. The probability of generating current observation symbol depends only on current state. That is $P(O \mid Q, \lambda) = \prod_{t=1}^{T} P(o_t \mid q_t, \lambda)$ (2.1)

iii. Amount of rainfall is considered to be low if it is below 341mm

iv. Amount of rainfall is considered to be moderate if it is within the range (341- 680)mm

v. Amount of rainfall is considered to be high if it is above 680mm

The classification of states in this research work, is based on the growing seasons of the major crops in the study area and the data collected. Now, based on our study area, let the seasonal rainfall for crop production be modelled by a three state Hidden Markov model and six observations.

Table 2.1: Seasons and Crops being Grown

Seasons	Crops grown	Months within a season
Season 1	Maize, melon ,Ground nut	March-June
Season 2	Rice ,soya beans	July-September
Season 3	Cowpea, Melon	October-February

2.1.2 The Observations of the Hidden Markov Model

The possible observations within each of the seasons are

$A = O_1 =$ (Low rainfall well spread)

$B = O_2 =$ (Low rainfall not well spread)

$C = O_3 =$ (Moderate rainfall well spread)

$D = O_4 =$ (Moderate rainfall not well spread)

$E = O_5 =$ (High rainfall well spread)

$F = O_6 =$ (High rainfall not well spread)

Rainfall Well Spread: Means rainfall that spreads evenly across the growing season

Rainfall Not Well Spread: Means rainfall that is not spread evenly across the growing season

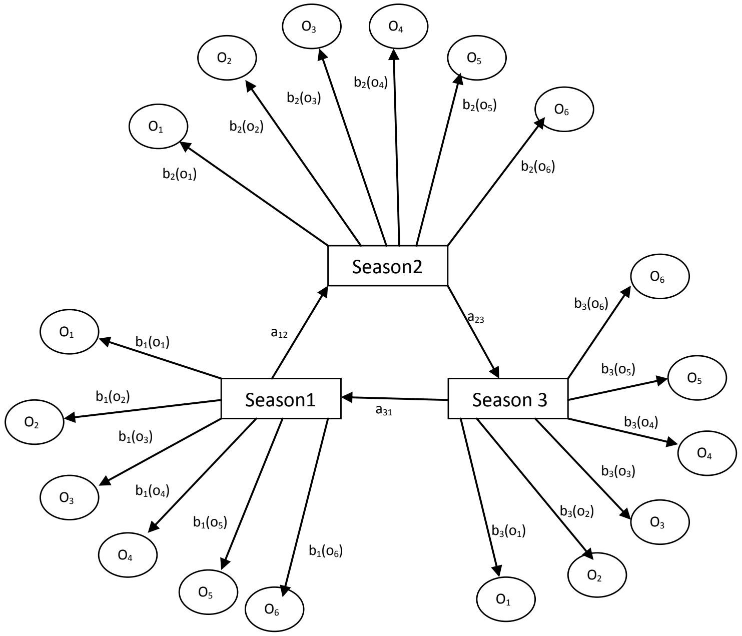

Figure 2.1 : The Transition Diagram of the Seasonal Rainfall Model

From Figure 2.1 we formulate the following

Season1 emissions

$b_1(o_1) = P(o_1\ at\ t \mid q_1\ at\ t)$, $b_1(o_2) = P(o_2\ at\ t \mid q_1\ at\ t)$, $b_1(o_3) = P(o_3\ at\ t \mid q_1\ at\ t)$

$b_1(o_4) = P(o_4\ at\ t \mid q_1\ at\ t)$, $b_1(o_5) = P(o_5\ at\ t \mid q_1\ at\ t)$,$b_1(o_6) = P(o_6\ at\ t \mid q_1\ at\ t)$

Season 2 emissions

$b_2(o_1) = P(o_1\ at\ t \mid q_2\ at\ t)$, $b_2(o_2) = P(o_2\ at\ t \mid q_2\ at\ t)$, $b_2(o_3) = P(o_3\ at\ t \mid q_2\ at\ t)$

$b_2(o_4) = P(o_4\ at\ t \mid q_2\ at\ t)$, $b_2(o_5) = P(o_5\ at\ t \mid q_2\ at\ t)$, $b_2(o_6) = P(o_6\ at\ t \mid q_2\ at\ t)$

Season 3 emissions

$b_3(o_1) = P(o_1\ at\ t \mid q_3\ at\ t)$, $b_3(o_2) = P(o_2\ at\ t \mid q_3\ at\ t)$, $b_3(o_3) = P(o_3\ at\ t \mid q_3\ at\ t)$

$b_3(o_4) = P(o_4\ at\ t \mid q_3\ at\ t)$, $b_3(o_5) = P(o_5\ at\ t \mid q_3\ at\ t)$, $b_3(o_6) = P(o_6\ at\ t \mid q_3\ at\ t)$

2.1.3Transition Probability Matrix for the states

The transitions between states are given by the matrix below

$$A = \begin{bmatrix} 0 & a_{12} & 0 \\ 0 & 0 & a_{23} \\ a_{31} & 0 & 0 \end{bmatrix} \tag{2.2}$$

2.1.4 Observation Probability Matrix

The observations emitted by each state are represented by the Matrix below

$$B = \begin{bmatrix} b_1(o_1) & b_1(o_2) & b_1(o_3) & b_1(o_4) & b_1(o_5) & b_1(o_6) \\ b_2(o_1) & b_2(o_2) & b_2(o_3) & b_2(o_4) & b_2(o_5) & b_2(o_6) \\ b_3(o_1) & b_3(o_2) & b_3(o_3) & b_3(o_4) & b_3(o_5) & b_3(o_6) \end{bmatrix} \tag{2.3}$$

Before training, the sum of column entries equal one, if there is observation otherwise zero

$$\sum_{k=1}^{3} b_k(o_m) = \begin{cases} 1 \\ 0 \; otherwise \end{cases} \tag{2.4}$$

After training, the sum of row entries must be one

$$\sum_{m=1}^{6} b_k(o_m) = 1 \tag{2.5}$$

2.1.5 Initial Probability Distribution

The initial probability distribution for the model is given below

$$\pi = [\pi_1 \quad \pi_2 \quad \pi_3] \tag{2.6}$$

Equation (2.6) is calculated using equation (2.7)

$$\pi_i = \frac{c_i}{L} \qquad i = 1, 2, 3 \tag{2.7}$$

Where c_i is the count of each state in the sequence of the states in the dataset

That is $c_1 + c_2 + c_3 = L$

2.1.6 The Hidden Markov Model

The overall hidden Markov model for Seasonal Rainfall Modelling for Crop Production is given by the compact notation below

$$\lambda = (A, B \;, \pi) \tag{2.8}$$

Hidden Markov Training

The hidden Markov model for Seasonal Rainfall Modelling for Crop Production developed shall be trained, using Baum Welch Algorithm mentioned in section

1.2.3.3 This will enable the model understand the historical data. At the end of the training, the new hidden Markov model λ^* will best fit the observed data. The new model will then make prediction with high precision.

2.1.8 The Application of the Seasonal Rainfall Model for Crop Production

We shall at this point provide numerical illustration for the model discussed in section 2.1. The data used in this research work, was collected from the archive of branch office of Nigerian Meteorological Agency, located in Minna Air port, Maikunkele. Maikunkele is the headquarters of Bosso Local Government Area of Niger state, located in north central Nigeria. It lies in 9.6490° N, 6.4530° E. The summary of the data is presented in Table 2.2 below

Table 2.2:A Summary of Seasonal Rainfall in Maikunkele Between 1980-2015 With Each Seasonal Observations

Year	1980			1981			1982			1983			1984		
State	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Observation	D	C	B	B	C	B	C	F	B	B	F	B	B	D	B

Year	1985			1986			1987			1988			1989		
State	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Observation	D	D	B	B	E	B	A	D	B	B	F	B	D	C	B

Year		1990			1991			1992			1993			1994	
State	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
observation	D	C	B	D	D	B	D	F	B	D	D	B	D	E	B

Year		1995			1996			1997			1998			1999	
State	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
observation	D	F	B	D	E	F	D	C	B	D	C	B	B	E	B

Year		2000			2001			2002			2003			2004	
State	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
observation	B	F	B	D	E	B	D	D	B	B	D	B	D	C	B

Year		2005			2006			2007			2008			2009	
State	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
observation	D	D	B	B	F	B	D	E	B	B	E	B	B	F	B

Year	2010			2011			2012			2013			2014		
State	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
observation	B	E	B	B	D	B	B	E	B	D	D	B	D	E	B

Year	2015		
State	1	2	3
observation	D	D	B

1 = Season **1**, **2** = Season2 and **3** = Season3

A = (Low rainfall well spread)

B = (Low rainfall not well spread)

C = (Moderate rainfall well spread)

D = (Moderate rainfall not well spread)

E = (High rainfall well spread)

F = (High rainfall not well spread)

2.1.8 Making Predictions With the Model

Making prediction for the three seasons is done along with the training of parameters of the model. The parameters of the model are initialized then trained using Baum-Welch algorithm in section 1.2.3.3 to attend Maximum likelihood. The forward probability of the training observation sequence is calculated from time t=1 to time T using Forward Algorithm (Rabiner, 1989). 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 seasons (season 2 at time T+2, and season 3 at time T+3). At the end of each year when the rainfall data recorded by the Nigerian Meteorological Agency is made available to the public, the predictions are compared with actual states and observations. If the prediction is not 100% accurate, it is included it in the training system to absorb the newest information. When the training is done and the parameters of the HMM are updated, we used the updated model to perform the predictions for the next three seasons using the previous method. To this end, we have developed two hidden Markov models, one of the models is to test for the reliability of the model and the other model is for future predictions. To avoid underflow of the forward algorithm we let the coefficient

$$c_t = \frac{1}{\sum_{t=1}^{N} \alpha_t(i)} \tag{2.9}$$

thus the new scaled value for α becomes

$$\hat{\alpha}_t(i) = c_t \times \alpha_t(i) = \frac{\alpha_t(i)}{\sum_{t=1}^{N} \alpha_t(i)} \tag{2.10}$$

2.1.10 Hidden Markov Model (HMM1) to Make Predictions for 2015 Seasons

HMM1 is a test HMM, to test for the reliability of the Model, the parameters of the HMM1 were estimated using the rainfall data from January, 1980 to February 2015, then made predictions for Season 1 (Mar-Jun, 2015), Season 2 (Jul-Sep, 2015) and Season 3 (Oct,2015 –Feb, 2016).

Transition Count Matrix

$$T = \begin{bmatrix} 0 & 35 & 0 \\ 0 & 0 & 35 \\ 34 & 0 & 0 \end{bmatrix} \tag{2.11}$$

Adding Pseudocount

All transitions and emissions never observed in the dataset are usually set to zero, as it can be seen in equation (2.11). This can be a problem during training of the HMM. A simple solution, is to add a constant called pseudocount to all the entries of the transition or emission matrix. If the constant is one, it is called laplace rule (Srinivas, 2006) and (Creighton and Hanash, 2003). This is illustrated in equation (2.12) below.

Pseudocount Transition Matrix

$$C = \begin{bmatrix} 1 & 36 & 1 \\ 1 & 1 & 36 \\ 35 & 1 & 1 \end{bmatrix} \tag{2.12}$$

The Transition Probability Matrix, Observation Count Matrix and Pseudocount Observation Matrix are given in Equations (2.13),(2.14) and(2.15) respectively.

$$A = \begin{bmatrix} 0.0263 & 0.9474 & 0.0263 \\ 0.0263 & 0.026 & 0.9474 \\ 0.9459 & 0.0270 & 0.0270 \end{bmatrix} \tag{2.13}$$

$$E = \begin{bmatrix} 1 & 14 & 1 & 19 & 0 & 0 \\ 0 & 0 & 7 & 10 & 10 & 8 \\ 0 & 35 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.14}$$

$$D = \begin{bmatrix} 2 & 15 & 2 & 20 & 1 & 1 \\ 1 & 1 & 8 & 11 & 11 & 9 \\ 1 & 36 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{2.15}$$

While Observation probability Matrix and Initial State Probabilities are given in Equations (2.16) and (2.17)

$$B = \begin{bmatrix} 0.5 & 0.2885 & 0.1818 & 0.6250 & 0.0769 & 0.0909 \\ 0.25 & 0.0192 & 0.7273 & 0.3438 & 0.8462 & 0.8182 \\ 0.25 & 0.6923 & 0.0909 & 0.0313 & 0.0769 & 0.0909 \end{bmatrix} \tag{2.16}$$

$$\pi = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \tag{2.17}$$

The overall HMM1 is represented by Equation (2.18)

$$\lambda_1 = (A, B, \pi) \tag{2.18}$$

After 250 iterations of the Baum Welch Algorithm, equation (2.18) stabilised to equation (2.19)

$$\lambda_1^* = (\hat{A}, \hat{B}, \hat{\pi}) \tag{2.19}$$

where

$$\hat{A} = \begin{bmatrix} 0.0278 & 0.9722 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 \\ 1.0000 & 0.0000 & 0.0000 \end{bmatrix} \tag{2.20}$$

$$\hat{B} = \begin{bmatrix} 0.0286 & 0.4000 & 0.0286 & 0.5429 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.1714 & 0.2857 & 0.3143 & 0.2286 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix} \tag{2.21}$$

and

$$\hat{\pi} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \tag{2.22}$$

Prediction for season 1, 2015

From our dataset (Table 2.2), the process is in Season3 at time T(Oct, 2014 to Feb, 2015) emitting Observation B. To get next observation at time T+1, we calculate the forward probability for BA, BB, BC, BD, BE, and BF using equation (1.25) and take the one with highest likelihood value in Table 2.3

Table 2.3: Likelihood Based Prediction Table for Season1, 2015.

	BA	BB	BC	BD	BE	BF
STATE 1	0.0286	0.4	0.0286	0.5429	0.0	0.0
STATE 2	0.0	0.0	0.0	0.0	0.0	0.0
STATE 3	0.0	0.0	0.0	0.0	0.0	0.0

From table 2.3, State 1 under BD, has the highest likelihood value, so is taken as most probable observation sequence at T+ 1 (that is, State 1 emitting observation D at T+1)

States: **3**(Oct, 2014 to Feb, 2015) → **1**(Mar –Jun, 2015)

↓ ↓

Observations: B D

Prediction for season 2, 2015

From the computation of Table 2.3, the process is in state1 at tme T+1(Mar-Ju, 2015) with observation D and sequence BD. To get the next sequence at time T+2, we calculate the forward probability for BDA, BDB, BDC, BDD, BDE and BDF and take the one with the highest likelihood value in Table 2.4

Table 2.4 : Likelihood Based Prediction Table for Season 2, 2015.

	BDA	BDB	BDC	BDD	BDE	BDF
STATE1	0.0008	0.0111	0.0008	0.0151	0	0
STATE2	0	0	0.1666	0.2778	0.3056	0.2222
STATE3	0	0	0	0	0	0

From Table 2.4, State 2 under BDE, has the highest likelihood value, so is taken as most probable observation sequence at T+ 2 (that is, State 2 emitting observation E at T+2).

States: **3**(Oct, 2014 to Feb, 2015) → **1**(Mar –Jun, 2015) → 2(Jul-sep, 2015)

↓ ↓ ↓

Observations: B D E

Prediction for Season 3, 2015

From the computation of Table 2.4, the process is in state 2 at time T+2 (Jul-Sep, 2015) with observation E and sequence BDE. To get the next sequence at time T+3, we calculate the forward probability for BDEA, BDEB, BDEC, BDED, BDEE and BDEF and take the one with the highest likelihood value in Table 2.5

Table 2.5 : Likelihood Based Prediction Table for Season 3, 2015

	BDEA	BDEB	BDEC	BDED	BDEE	BDEF
STATE1	0.0	0.0	0.0	0.0	0.0	0.0
STATE2	0.0	0.0	0.0	0.0	0.0	0.0
STATE3	0.0	1.0	0.0	0.0	0.0	0.0

From Table 2.5, State 3 under BDEB has the highest likelihood value, so is taken as most probable observation sequence at T+3, (that is, State 3 emitting observation B at T+3)

States: **3**(Oct, 2014 to Feb, 2015) → **1**(Mar –Jun, 2015)→ 2(Jul-Sep, 2015)→3(Oct,2015 –Feb, 2016)

↓ ↓ ↓ ↓

Observations: B D E B

Comparison of the Predicted states and Observations, and the Actual States and Observations from the Dataset

Predicted States and Observations for 2015 Seasons Using HMM1

States: **3**(Oct, 2014 to Feb, 2015)→ **1**(Mar –Jun, 2015)→2(Jul-Sep, 2015)→3(Oct,2015 –Feb, 2016)

↓ ↓ ↓ ↓

Observations: B D E B

Actual states and Observations from the Dataset

States:**3**(Oct, 2014 to Feb, 2015) → **1**(Mar –Jun, 2015) → 2(Jul-Sep, 2015) →3(Oct,2015 –Feb, 2016)

	↓	↓	↓	↓
Observations:	B	D	D	B

Hidden Markov Model(HMM2) to Make Predictions for 2016 Seasons

HMM2 was developed to make predictions for future seasons, the parameters of the HMM2 was estimated using Table 2.2, then made predictions for season 1(Mar-Jun, 2016), season 2(Jul-Sep, 2016) and season 3(Oct,2016 –Feb, 2017).

The Transition Count Matrix, Pseudocount transition matrix and the Transition Probability Matrix are given in Equations (2.23), (2.24) and (2.25) respectively

$$C = \begin{bmatrix} 0 & 36 & 0 \\ 0 & 0 & 36 \\ 35 & 0 & 0 \end{bmatrix} \tag{2.23}$$

$$T = \begin{bmatrix} 1 & 37 & 1 \\ 1 & 1 & 37 \\ 36 & 1 & 1 \end{bmatrix} \tag{2.24}$$

$$A = \begin{bmatrix} 0.0256 & 0.9487 & 0.0256 \\ 0.0256 & 0.0256 & 0.9487 \\ 0.9474 & 0.0263 & 0.0263 \end{bmatrix} \tag{2.25}$$

While Observation Count Matrix, Pseudocount Observation Matrix and Observation Probability Matrix are given in Equations (2.26), (2.27) and (2.28) respectively

$$E = \begin{bmatrix} 1 & 14 & 1 & 20 & 0 & 0 \\ 0 & 0 & 7 & 11 & 10 & 8 \\ 0 & 36 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.26}$$

$$D = \begin{bmatrix} 2 & 15 & 2 & 21 & 1 & 1 \\ 1 & 1 & 8 & 12 & 11 & 9 \\ 1 & 36 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{2.27}$$

$$B = \begin{bmatrix} 0.5 & 0.2885 & 0.1818 & 0.6176 & 0.0769 & 0.0909 \\ 0.25 & 0.0192 & 0.7273 & 0.3529 & 0.8462 & 0.8182 \\ 0.25 & 0.6923 & 0.0909 & 0.0294 & 0.0769 & 0.0909 \end{bmatrix} \tag{2.28}$$

The Initial probabilities is represented by Equation (2.29)

$$\pi = [1 \quad 0 \quad 0] \tag{2.29}$$

The Overall HMM2 is represented by Equation (2.30)

$$\lambda_2 = (\pi, A, B) \tag{2.30}$$

After 250 iterations of the Baum Welch Algorithm, equation (2.30) stabilised to equation (2.31)

$$\lambda_2^* = (\hat{A}, \hat{B}, \hat{\pi}) \tag{2.31}$$

where

$$\hat{A} = \begin{bmatrix} 0.0271 & 0.9729 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 \\ 1.0000 & 0.0000 & 0.0000 \end{bmatrix} \tag{2.32}$$

$$\hat{B} = \begin{bmatrix} 0.0278 & 0.3881 & 0.0278 & 0.5556 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.1667 & 0.3055 & 0.3056 & 0.2222 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix} \tag{2.33}$$

and

$$\hat{\pi} = [1 \quad O \quad O] \tag{2.34}$$

Prediction for Season 1, 2016

From our dataset Table 2.2, the process is in Season3 at time T(Oct, 2015 to Feb, 2016) emitting Observation B. Now, to get the next emission at time T+1 (Mar –Jun, 2016) we calculate the forward probability for BA, BB, BC, BD, BE, and BF using equation (1.25) and take the one with highest likelihood value in Table 2.6

Table 2.6: Likelihood Based Prediction Table for Season1, 2016.

	BA	BB	BC	BD	BE	BF
STATE1	0.0278	0.3887	0.0278	0.5553	0	0
STATE2	0	0	0.0008	0.0001	0.0001	0.0001
STATE3	0	0	0	0	0	0

From Table 2.6, State1 under BD, has the highest likelihood value, so is taken as most probable observation sequence at T+1, (that is, State1 emitting observation D at T+1).

States: **3**(Oct, 2015 to Feb, 2016) $\rightarrow$ **1**(Mar –Jun, 2016)

$\downarrow$ $\downarrow$

Observations: B D

Prediction for Season 2, 2016

From the computation of Table 2.6, the process is in state1 at time T+1(Mar-Jun, 2016) with observation D and sequence BD. To get the next sequence at time T+2(Jul-sep, 2016) we calculate the forward probability for BDA, BDB, BDC, BDD, BDE and BDF and take the one with the highest likelihood value in Table 2.7

Table 2.7: Likelihood Based Prediction Table for Season 2, 2016.

	BDA	BDB	BDC	BDD	BDE	BDF
STATE1	0.0008	0.0105	0.0008	0.0151	0	0
STATE2	0	0	0.1621	0.2971	0.2972	0.2161
STATE3	0	0.0003	0	0	0	0

From Table 2.7, State 2 under BDE, has the highest likelihood value, so is taken as most probable observation sequence at T+2 (that is, State 2 emitting observation E at T+2)

$$
\begin{array}{lccccc}
\text{States:} & \mathbf{3}(\text{Oct, 2015 to Feb, 2016}) & \rightarrow & \mathbf{1}(\text{Mar –Jun, 2016}) & \rightarrow & 2(\text{Jul-sep, 2016}) \\
 & \downarrow & & \downarrow & & \downarrow \\
\text{Observations:} & B & & D & & E
\end{array}
$$

Prediction for Season 3, 2016

From the computation of Table 2.7, the process is in state 2 at time T+2(Jul-Sep, 2016) with observation E and sequence BDE. To get the next sequence at time T+3(Oct,2016 –Feb, 2017) we calculate the forward probability for BDEA, BDEB, BDEC, BDED, BDEE and BDEF and take the one with the highest likelihood value in Table 2.8

Table 2.8: Likelihood Based Prediction Table for Season3, 2016.

	BDEA	BDEB	BDEC	BDED	BDEE	BDEF
STATE 1	0.0	0.0	0.0	0.0	0.0	0.0
STATE 2	0.0	0.0	0.0	0.0	0.0	0.0
STATE 3	0.0	1.0	0.0	0.0	0.0	0.0

From Table 2.8, State 3 under BDEB has the highest likelihood value, so is taken as most probable observation sequence at T+3, (that is, State 3 emitting observation b at T+3)

Predicted states and Observations for 2016 seasons Using HMM2

States: **3**(Oct, 2015 to Feb, 2016)→ **1**(Mar –Jun, 2016) → **2**(Jul-Sep, 2016) → **3**(Oct,2016 –Feb, 2017)

	↓	↓	↓	↓
Observations:	B	D	E	B

The parameters of the HMM1 were estimated using rainfall data from January, 1980 to February 2015, after 250 iterations of the Baum Welch algorithm, λ_1 stabilised to a new model, λ_1*, this new model was then used to make predictions for 2015 seasons. As shown, HMM1 was in State 3 at time T(Oct. 2014 to Feb. 2015) emitted observation B then made transition to season1 at time T+1 (Mar – Jun, 2015) governed by first order Markov dependence, emitting observation D. Similar interpretation is given to transition to season 2 at T+2(Jul-Sep., 2015) and transition to season 3 at time T+3 (Oct. 2015 –Feb. 2016) emitting observation E and B respectively. HMM1 has shown a 100% accuracy in States predictions and 75% in observations predictions when compared with the actual states and observations from the dataset.

The parameters of the HMM2 were estimated using rainfall data in Table 2.2, after 250 iterations the Baum Welch algorithm, λ_2 stabilised to a new model, λ_2*, this new model was then used to make predictions for 2016 seasons. It was found that HMM2 was in State 3 at time T(Oct. 2015 to Feb. 2016) emitted observation B then made transition to season1 at time T+1 (Mar – Jun, 2016) governed by first order Markov dependence, emitting observation D. Similar interpretation is given to transition to season2 at T+2(Jul-Sep., 2016) and transition to Season 3 at time T+3 (Oct. 2016 – Feb. 2017) emitting observation E and B respectively.

2.2 Annual Rainfall Pattern for Crop Production in Jos Plateau State, Nigeria

Plateau state is one of the north central states. It has an estimated population of about three million people, with an area of about 26,899 square kilometres. The state is named after the picturesque Jos Plateau, a mountainous area in the north of the state with captivating rock formations. It is located between latitude 80°24'N and longitude 80°32'E. The altitude ranges from around 1,200 meters (about 4000 feet) to a peak of 1,829 metres above sea level in the Shere Hills range near Jos (Plateau State, 2016).

In each year, rainfall onset and recession have become a very big problem to the farmers and policy makers in planning for crop cultivation. This is because rainfall starts early or late and stops early or late in a year. This had over the years affected crop cultivation and planting and it had consequently led to poor harvest. the Hidden Markov Model is developed on basis of observed data in the study area and it is capable of predicting rainfall onset, recession, spread and amount within a year.

Formulation of the Model

In this model, amount of rainfall is considered as state of the Hidden Markov Model while onset, recession, distribution (spread) of the rainfall within a year is taken as emission of the model. As a result, we make the following assumptions

i. The transition of annual rainfall state to another state in a year follows a Markov chain of first order dependence as represented by equation (1.1)
ii. The probability of generating current observation symbol depends only on current state, that is

$$P(O \mid Q, \lambda) = \prod_{t=1}^{T} P(o_t \mid q_t, \lambda) \tag{2.35}$$

iii. Rainfall is said to start early if it starts within (January-March)
iv. Rainfall is said to start late if it starts on the month of April or beyond
v. Rainfall is said to stop early if the rainfall does not exceed the month of October of a year
vi. Rainfall is said to stop late if the rainfall exceeds 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 ≤ 800mm)

State2: Moderate Rainfall (801mm≤ rainfall amount ≤1300mm)

State3: High Rainfall (Rainfall amount >1300mm)

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

Let

$A = O_1 =$ (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 2.2

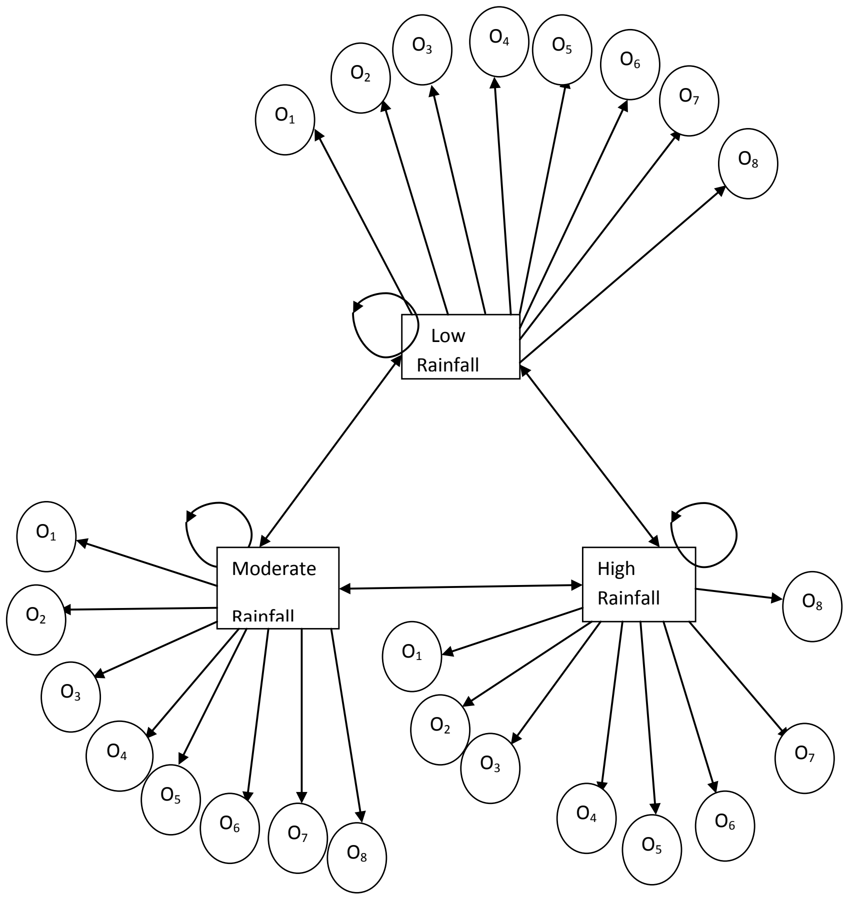

Figure 2.2: Transition Diagram of the Annual Rainfall Model

From Figure 3.1, we have the following emissions

State 1 Emission

$b_l(o_1) = P(o_1 \ at\ t \mid q_l \ at\ t)$, $b_l(o_2) = P(o_2 \ at\ t \mid q_l \ at\ t)$, $b_l(o_3) = P(o_3 \ at\ t \mid q_l \ at\ t)$

$b_l(o_4) = P(o_4 \ at\ t \mid q_l \ at\ t)$, $b_l(o_5) = P(o_5 \ at\ t \mid q_l \ at\ t)$, $b_l(o_6) = P(o_6 \ at\ t \mid q_l \ at\ t)$

$b_l(o_7) = P(o_7 \ at\ t \mid q_l \ at\ t)$, $b_l(o_8) = P(o_8 \ at\ t \mid q_l \ at\ t)$

State 2 Emissions

$b_m(o_1) = P(o_1 \ at\ t \mid q_m \ at\ t)$, $b_m(o_2) = P(o_2 \ at\ t \mid q_m \ at\ t)$, $b_m(o_3) = P(o_3 \ at\ t \mid q_m \ at\ t)$

$b_m(o_4) = P(o_4 \ at\ t \mid q_m \ at\ t)$, $b_m(o_5) = P(o_5 \ at\ t \mid q_m \ at\ t)$, $b_m(o_6) = P(o_6 \ at\ t \mid q_m \ at\ t)$

$b_m(o_7) = P(o_7 \ at\ t \mid q_m \ at\ t)$, $b_m(o_8) = P(o_8 \ at\ t \mid q_m \ at\ t)$

State 3 Emissions

$b_h(o_1) = P(o_1 \ at\ t \mid q_h \ at\ t)$, $b_h(o_2) = P(o_2 \ at\ t \mid q_h \ at\ t)$, $b_h(o_3) = P(o_3 \ at\ t \mid q_h \ at\ t)$

$b_h(o_4) = P(o_4 \ at\ t \mid q_h \ at\ t)$, $b_h(o_5) = P(o_5 \ at\ t \mid q_h \ at\ t)$, $b_h(o_6) = P(o_6 \ at\ t \mid q_h \ at\ t)$

$b_h(o_7) = P(o_7 \ at\ t \mid q_h \ at\ t)$, $b_h(o_8) = P(o_8 \ at\ t \mid q_h \ at\ t)$

2.2.2 Transition Probability Matrix for The states

The transitions between the states are represented by the Matrix A

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{2.36}$$

2.2.3 Observation Probability Matrix

The observations emitted by each state are represented by the Matrix B

$$B = \begin{bmatrix} b_l(o_1) & b_l(o_2) & b_l(o_3) & b_l(o_4) & b_l(o_5) & b_l(o_6) & b_l(o_7) & b_l(o_8) \\ b_m(o_1) & b_m(o_2) & b_m(o_3) & b_m(o_4) & b_m(o_5) & b_m(o_6) & b_m(o_7) & b_m(o_8) \\ b_h(o_1) & b_h(o_2) & b_h(o_3) & b_h(o_4) & b_h(o_5) & b_h(o_6) & b_h(o_7) & b_h(o_8) \end{bmatrix} \tag{2.37}$$

2.2.4 Initial Probability Distribution

The initial probability distribution of the model is given by Equation (2.39)

$$\pi = [\pi_1 \quad \pi_2 \quad \pi_3] \tag{2.39}$$

2.2.5 Application of Annual Rainfall Pattern for Crop Production in Jos, Plateau State, Nigeria

The data used in this research work, was 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. It is located between latitude 9.9503^0N and longitude 8.8828^0E. The summary is presented in Table 2.9.

Table 2.9: A summary of Annual Rainfall in Jos between1977-2015 with States and their Observations

Year	State	Observation
1977	1	B
1978	2	A
1979	2	A
1980	2	H
1981	3	G
1982	2	B
1983	2	B
1984	3	B
1985	2	B
1986	3	B
1987	2	B
1988	2	B
1989	3	B
1990	2	H
1991	3	B
1992	3	D
1993	3	H

1994	3	H
1995	3	H
1996	3	B
1997	2	D
1998	2	H
1999	3	B
2000	2	B
2001	1	H
2002	2	B
2003	1	H
2004	3	B
2005	2	H
2006	2	B
2007	1	H
2008	2	H
2009	2	H
2010	1	H
2011	2	H
2012	2	H
2013	3	H
2014	3	B
2015	3	B

2.2.6 Testing the Reliability of the Model

In order to test for the reliability of the model, we divided the dataset into two sets, one training set and the other test set. We estimated the parameter of the test HMM1 using the rainfall data from 1977-1994, then used it to predict annual rainfall for 1995, 1996 and 1997. The Transition Count Matrix, Transition Probability Matrix and Observation Count Matrix are given in Equations (2.39), (2.40) and (2.41) respectively.

$$C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 5 & 5 \\ 0 & 4 & 3 \end{bmatrix} \tag{2.39}$$

$$A = \begin{bmatrix} 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.5000 & 0.5000 \\ 0.0000 & 0.5714 & 0.4286 \end{bmatrix} \tag{2.40}$$

$$E = \begin{bmatrix} 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 3 & 0 & 1 & 0 & 0 & 1 & 2 \end{bmatrix} \tag{2.41}$$

While Observation probability Matrix and Initial State Probabilities are given in Equations (2.42) and (2.43) respectively

$$B = \begin{bmatrix} 0.0000 & 0.2000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 1.0000 & 0.5000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.5000 \\ 0.0000 & 0.3000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 1.0000 & 0.5000 \end{bmatrix} \tag{2.42}$$

$$\pi = [0.0556 \quad 0.5000 \quad 0.4444] \tag{2.43}$$

The overall HMM1 is represented by Equation (2.44)

$$\lambda_1 = (A, B, \pi) \tag{2.44}$$

After 500 iterations of the Baum Welch Algorithm using Matlab, equation (2.44) stabilised to equation (2.45)

$$\lambda_1^* = (\hat{A}, \hat{B}, \hat{\pi}) \tag{2.45}$$

where

$$\hat{A} = \begin{bmatrix} 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.5000 & 0.5000 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \tag{2.46}$$

$$\hat{B} = \begin{bmatrix} 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.9999 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0001 \\ 0.0000 & 0.6000 & 0.0000 & 0.0667 & 0.0000 & 0.0000 & 0.0667 & 0.2667 \end{bmatrix} \tag{2.47}$$

and

$$\hat{\pi} = [\mathrm{O} \quad \mathrm{O} \quad 1] \tag{2.48}$$

Prediction for 1995

From the dataset (Table 2.9), the process is in state 3 at time T(1994) emitting Observation H. To get next observation at time T+1, we calculate the forward probability for HA, HB, HC, HD, HE, HF, HG, HH using equation (1.25) and take the one with highest likelihood value.

Table 2.10: Likelihood Based Prediction Table for 1995

	HA	HB	HC	HD	HE	HF	HG	HH
STATE1	00	0.0	0.0	0.0	0.0	0.0	0.0	0.0
STATE2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
STATE3	0.0	0.6	0.0	0.0667	0.0	0.0	0.0667	0.2667

From Table 2.10, State 3 under HB has the highest likelihood value and is taken as most probable observation sequence at time T+1 (that is, State 3 emitting observation B at time T+1)

States **3**(1994) → **3**(1995)

↓ ↓

Observations H B

Prediction for 1996

From the computation of Table 2.10, the process is in state 3 at time T+1(1995) with observation B and sequence HB. To get the next sequence, we calculate the forward probability for HBA, HBB, HBC, HBD, HBE, HBF, HBG and HBH and take the one with the highest likelihood value in Table 2.11.

Table 2.11: Likelihood Based Prediction Table for 1996

	HBA	HBB	HBC	HBD	HBE	HBF	HBG	HBH
STATE1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
STATE2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
STATE3	0.0	0.6	0.0	0.0667	0.0	0.0	0.0667	0.2667

From Table 2.11, State 3 under HBB has the highest likelihood value, is taken as most probable observation sequence at time T+2, (that is, State 3 emitting observation B at T+2)

States 3(1994) → 3(1995) → 3(1996)

↓ ↓ ↓

Observations H B B

Prediction for 1997

From the computations of Table 2.11, the process is in state 3 at time T+2(1996) with observation B and sequence HBB. To get the next sequence, we calculate the forward probability for HBBA, HBBB, HBBC, HBBD, HBBE, HBBF, HBBG and HBBH and take the one with the highest likelihood value in table 2.12

Table 2.12: Likelihood Based Prediction Table for 1997

	HBBA	HBBB	HBBC	HBBD	HBBE	HBBF	HBBG	HBBH
STATE1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
STATE2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.001
STATE3	0.0	0.6	0.0	0.066	0.0	0.0	0.066	0.266

From Table 2.12, State 3 under HBBB has the highest likelihood value, is taken as most probable observation sequence at time T+3 (that is, State 3 emitting observation B at T+3)

States **3**(1994) → **3**(1995) → 3(1996) → 3(1997)

↓ ↓ ↓ ↓

Observations: H B B B

Comparison of the Predicted States and Observations, and the Actual States and Observations From the 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

2.2.7 Hidden Markov Model for Future Predictions (HMM2)

In order to make predictions for the future years, the whole dataset (rainfall data from 1977 to 2015) was used to estimate the parameters of the model, then make predictions for 2016, 2017 and 2018 .The Transition Count Matrix, Transition Probability Matrix and Observation Count Matrix are given in Equations (2.49), (2.50) and (2.51) respectively.

$$C = \begin{bmatrix} 0 & 4 & 1 \\ 4 & 8 & 7 \\ 0 & 7 & 7 \end{bmatrix} \tag{2.49}$$

$$A = \begin{bmatrix} 0.0000 & 0.8000 & 0.200 \\ 0.2100 & 0.4211 & 0.3684 \\ 0.0000 & 0.5000 & 0.5000 \end{bmatrix} \tag{2.50}$$

$$B = \begin{bmatrix} 0 & 2 & 0 & 0 & 0 & 0 & 0 & 4 \\ 2 & 8 & 0 & 1 & 0 & 0 & 0 & 8 \\ 0 & 8 & 0 & 1 & 0 & 0 & 1 & 4 \end{bmatrix} \tag{2.51}$$

While Observation Probability Matrix and Initial State Probabilities are given in Equations (2.52) and (2.53) respectively

$$B = \begin{bmatrix} 0.0000 & 0.1111 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.2500 \\ 1.0000 & 0.4444 & 0.0000 & 0.5000 & 0.0000 & 0.0000 & 0.0000 & 0.5000 \\ 0.0000 & 0.4444 & 0.0000 & 0.5000 & 0.0000 & 0.0000 & 1.0000 & 0.2500 \end{bmatrix} \tag{2.52}$$

$$\pi = [0.1282 \quad 0.4872 \quad 0.3846] \tag{2.53}$$

The overall HMM2 is represented by Equation (2.54)

$$\lambda_2 = (A, B, \pi) \tag{2.54}$$

After 500 iterations of the Baum Welch Algorithm, equation (2.54) stabilised to equation (2.55)

$$\lambda_2^* = (\hat{A}, \hat{B}, \hat{\pi}) \tag{2.55}$$

where

$$\hat{A} = \begin{bmatrix} 0 & 0.6332 & 0.3668 \\ 0.7668 & 0.2332 & 0.0000 \\ 0 & 0.3643 & 0.6357 \end{bmatrix} \tag{2.56}$$

$$\hat{B} = \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \\ 0.1338 & 0.4137 & 0.0000 & 0.1338 & 0.0000 & 0.0000 & 0.0000 & 0.3187 \\ 0.0000 & 0.9220 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0780 & 0.0000 \end{bmatrix} \tag{2.57}$$

and

$$\hat{\pi} = [0 \quad 0 \quad 1] \tag{2.58}$$

Prediction for 2016

From the dataset(Table 2.9), the process is in state 3 at time T(2015) emitting Observation B. To get next observation at time T+1, we calculate the forward probability for BA, BB, BC, BD, BE, BF, BG, BH using equation (1.25) and take the one with highest value in Table 2.13.

Table 2.13: Likelihood Based Prediction Table for 2016

	BA	BB	BC	BD	BE	BF	BG	BH
STATE1	0	0	0	0	0	0	0	0.2213
STATE2	0.0437	0.1351	0	0.0427	0	0	0	0.1041
STATE3	0	0.4170	0	0	0	0	0.0353	0

From Table 2.13, State3 under BB has the highest likelihood value and is taken as most probable observation sequence at T+1 (that is, State3 emitting observation B at T+1)

States: 3(2015) $\rightarrow$ 3(2016)

$\downarrow$ $\downarrow$

Observation B B

Prediction for 2017

From the computations of Table 2.13, the process is in state 3 at time T+1(2016) with observation B and sequence BB. To get the next sequence at time T+2, we calculate the forward probability for BBA, BBB, BBC, BBD, BBE, BBF, BBG and BBH and take the one with the highest likelihood value in Table 2.14

Table 2.14: Likelihood Based Prediction Table for 2017

	BBA	BBB	BBC	BBD	BBE	BBF	BBG	BBH
STATE1	0	0	0	0	0	0	0	0.1876
STATE2	0.0445	0.1375	0	0.0437	0	0	0	0.1041
STATE3	0	0.4427	0	0	0	0	0.0375	0

From table 2.14, State 3 under BBB has the highest likelihood value and is taken as most probable observation sequence at T+2 (that is, State 2 emitting observation B at T+2)

$$
\begin{array}{llll}
\text{States:} & \mathbf{3}(2015) \rightarrow & \mathbf{3}(2016) \rightarrow & 3(2017) \\
& \downarrow & \downarrow & \downarrow \\
\text{Observations:} & B & B & B
\end{array}
$$

Prediction for 2018

From the computations of Table 2.14, the process is in state 3 at time T+2(2017) with observation B and sequence BBB. To get the next sequence at time T+3, we calculate the forward probability for BBBA, BBBB, BBBC, BBBD, BBBE, BBBF, BBBG and BBBH and take the one with the highest likelihood value in Table 2.15.

Table 2.15: Likelihood Based Prediction Table for 2018

	HBBA	HBBB	HBBC	HBBD	HBBE	HBBF	HBBG	HBBH
STATE1	0	0	0	0	0	0	0	0.1876
STATE2	0.0442	0.1367	0	0.0442	0	0	0	0.1053
STATE3	0	0.4413	0	0	0	0	0.0375	0

From Table 2.15, State 3 under HBBB has the highest likelihood value, is taken as most probable observation sequence at time T+3 (that is, State 3 emitting observation B at T+3)

Predicted states and Observations Using HMM2

States: **3**(2015) → **3**(2016) → 2(2017) →1(2018)

↓ ↓ ↓ ↓

Observations: B B B B

The parameters of the HMM1 were estimated using rainfall data from 1977-1994, after 500 iterations of the Baum algorithm, λ_1 stabilised to a new model, λ_1^*, this new model was then used to make predictions from 1994-1997. 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). This model has 75% accuracy is state and 50% in observation in the predictions, this model was purposely developed to test for the reliability of the model.

The parameters of the HMM2 were estimated using the whole dataset (rainfall data from 1977-2015), after 500 iterations of the Baum Welch algorithm, λ_2 stabilised to a new model , λ_2^*. This model was used to make predictions for annual rainfall pattern for the future years(2016, 2017 and 2018). The prediction 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 made transition to state 2 (moderate rainfall) at time T+2(2017) and later to state1(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).

REFERENCES

Abubakar, U. Y., Lawal, A., & Muhammed, A. (2013). The Use of Markov Model in Continuous Time for Prediction of Rainfall for Crop Production. *IOSR Journal of Mathematics,* 7(1)*,* 38-45.

Adams, R.M., Hurd, B.H., Lenhart, S., & Leary, N. (1998). Effects of global climate Change On Agriculture: An interpretative review, *Climate Change*, 11, 19–30.

Ahn, P.M. (1993). Tropical Soils and Fertiliser Use. London: Longman Scientific Technical.

Andrew, W., Robertson, S .K., & Padhraic, S. (2003). Hidden Markov models For modelling daily rainfall occurrence over Brazil. *Technical Report UCI-ICS 03-27* Information and Computer Science University of California, Irvine.

Andrew W. R., Amor V. M., & James W. H. (2007). Downscaling of Seasonal Precipitation for Crop Simulation. *Journal of Applied Meteorology and Climatology.* 46, 677-693

Arumugam, P., & Karthik, S.M. (2016).Stochastic Modelling in Yearly Rainfall at Tirunelveli District,Tamil Nadu, India. *International Conference on Processing of Materials, Minerals and Energy* (July 29th – 30th) 2016, Ongole, Andhra Pradesh, India.

Barkotulla, .M. A. B .(2010). Stochastic Generation of the Occurrence and Amount of Daily Rainfall. *Pakistan journal of statistics and operation research*. *6 (1),*61-63.

Baldi, P., & Brunak, S. (2001). Bioinformatics. Cambridge: A Bradford Book, The MIT Press,.

Blessing Smart.(2015). Top 6 Geopolitical zones in Nigeria. Information guide in Nigeria. Retrieved from http://infoguidenigeria.com/top-6-geopolitical-zones-nigeria.

Chulsang,Y., Jinwook, L., & Yonghun, R. (2016).Markov Chain Decomposition of Monthly Rainfall into Daily Rainfall. *Evaluation of Climate Change Impact.* Hindawi Publishing Corporation Advances in Meteorology Volume 2016, 1-10.

Chaouche, A., & Parent, E. (1999). Bayesian Identification and Validation of a daily rainfall model under monsoon conditions. *Hydrology. Science Journal,* 44(2), 199-220.

Cox, D.R., & Isham, V. (1994). Stochastic models of precipitation: Statistics for the environment 2: *Water issues*. Wiley, New York, 3-18.

Creighton .C., & Hanash. S. (2003). Mining gene expression databases for association rules. *Bioinformatics*, 19,79–86.

Danilo, B .(2014). Hidden Markov Models: An Approach to Sequence Analysis In Population Studies, *National Center of Competence in Research LIVES Institute for Demographic and Life Course Studies,* University of Geneva, Switzerland, PPA 2014 Annual Meeting.

Delleur, J.W.,Chang,T. J., & Kavvas, M. L. (1989). Simulation Models of Sequences of wet and dry days. *Journal of Irrigation and Drainage Engineering,* 115(3), 344-357.

Durbin, R., Eddy, S.Y., Krogh. A., & Mitchison .G.(1998). Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. Cambridge, UK: Cambridge University Press.

Enza, M., & Daniele, T.(2007). Hidden Markov models for scenario generation. *Journal of Management Mathematics*, 19, 379-401.

Fink,G.A.(1989). Markov Models for Pattern Recognition. From Theory to Applications, Second Edition. London, UK: Springer.

Fisher, R. (1924). The Influence of the Rainfall on the yield of wheat a Rothamsted. Philosophical Transaction of the Royal Societies of London, 2, 120-123.

Gabriel, K.R., & Neumann, J. (1962). A Markov chain Model for Daily Rainfall Occurrences at Tel Aviv. *Quart Journal Ray Meteorology society,* 88, 90-95.

Geng, S., Penning. D, Vries F.W.T., & Supit, I. (1986). A Simple Method for Generating Daily Rainfall Data. *Agricultural. Forest Meteorol*ogy, 36, 363–376.

Hay, L. E., McCabe, J. G. J., Wolock, D. M.,& Ayers, M. A. (1991). Simulation of Precipitation by weather type analysis. Water Resources Research, 27, 493-501.

Hyang-Won. L(2014). Hidden Markov Models, Probability and Statistics Spring,

P1-18

Hirsch H.G.(2001).HMM Adaptation for Applications in Telecommunication. Speech Communication, 34(1-2), 127-139.

Hulme, M., Osborn, T., & John, T .(1998). Precipitation Sensitivity of Global Warming: Comparison of Observations with HADCM2 Simulations. *Geophysical Research Letter*, 25, 3379-3382.

Iloeje N.P. (1981). A New Geography of Nigeria. Nigeria: Longman Nigeria Limited.

Jae, Y. S. (2012). Investigation of Non-homogeneous Hidden Markov Model and Their Application to Spatially-Distributed Precipitation Types, a master of science Thesis Submitted to the Office of Graduate Studies of Texas A&M University, USA.

James, H., Matis, C., R., Perry, D.E., & Boudreaux D.J.A.(1989). Markov Chain Forecasts of Cotton Objective Yield. *United States National Agricultural Statistics. Service Research and Application Divisions, SRB Research Report* Number SRB-89-11.

Jimoh, O.D., & Webster, P.(1996). Optimum Order of Markov chain for daily rainfall in Nigeria *Journal of Hydrology*,185, 45-69.

Jimoh, O. D., & Webster, P. (1999). Stochastic Modelling Daily Rainfall in Nigeria: intra-annual variation of model parameters. *Journal of Hydrology*, 222, 1- 17.

Katrin, S., Sebastian, M., Gerd, S., & Christoph, K. (2016). Compound Extremes in a Changing Climate – A Markov Chain Approach. *Nonlinear Processes in Geophysics Discussions.*

Lawal, .A. (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.

Maheepala, S., & Perera, C.J.C.(1996).Monthly hydrologic data data generation by disaggregation. *Journal of Hydrology*. 178, 277-291.

Mouelhi, S., Nemri, S., Jebari, S., Slimani, M. (2016). Using the Markov Chain for the Generation of Monthly Rainfall Series in a Semi-Arid Zone. *Open Journal of Modern Hydrology,* 6, 51-65.

Parzen, E. (1962). Stochastic Processes, Second Edition. Oakland California, USA: Holden- Day Series in Probability and Statistics.

Plateau State. (2016, November 29). In Wikipedia, The Free Encyclopedia. Retrieved08:18,November29,2016,from https://en.wikipedia.org/w/index.php?title = Platea State & oldid=752062055.

Porter, J. W., & Pink, B. J. (1991). A Method of synthetic fragments for disaggregation in Stochastic data generation. *Hydrology and Water Resources Symposium*,187-191.

Przemyslaw D. (2011). Hidden Markov Models, Theory and Applications: Intech Crotia.

Rabiner, L.R.(1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE*, 77, 257-286.

Raheem, M. A., Yahya, W.B., & Obisesan, K.O. (2015). A Markov Chain Approach on Pattern of Rainfall Distribution. *Journal of Environmental Statistics, 7(1), 451-457.*

Ross S.M. (1989). Introduction to probability Models. London: Academic Press, Ltd.

Saran, S, T. K., Kitti .C., Preesan .R., & Thitiporn, C, Itsuo, K.(2014). Rice Cultivation an Harvest Date Estimation Using MODIS NDVI Time-series. *Data Nagatsuta-cho, Midori-ku, Yokohama Japan*, 226-8503.

Srinivas, A. (2006). Handbook of Computational Molecular Biology. Iowa state University, Ames, Iowa, USA. Chapman & Hall/CRC, Taylor & Francis Group.

Tamil,S., & Samuel, S. (2011). Stochastic Modelling of Annual Rainfall at Tamil Nadu. *Universal Journal of Environmental Research and Technology*, 1(4) 566-570.

Tang, X. (2004). Autoregressive Hidden Markov Model with Application in An EL NiNo Study. A masters Thesis Submitted to the College of Graduate Studies and Research Department of Mathematics and Statistics University of Saskatchewan, Saskatoon.

Thyer, M., & Kuzcera, G. (1999). Modelling long-term persistence in rainfall time series: Sydney rainfall case study. *Hydrology and Water Resources Symposium*, 550-555.

Vami H., N., Mahaman, B. S., Sorin. P., Romulus, T., Bogdan, B., Eric, V. D., Fernand, K., Monica, B., & Kouadio, A. (2014). Markovian Approach for analysis and Prediction of Monthly Precipitation Field in the Department of Sinfra .*International Journal of Engineering Research and General Science,* 2(1), 34-49.

Vaseghi, S.V.(2006). Advanced Digital Signal Processing and Noise Reduction. 4th Edition US: Wiley.

Vent, M .S., & Fritz, C.G.(1957).Weather Influence on the Size of US cord Soya beans. *Department of Agronomy, Agricultural Experimental Station Seasonal Journal*.

Wei .L., Tana, F. Y., & Zulkifli, Y.(2013). Non-Homogeneous Hidden Markov Model for Daily Rainfall Amount in Peninsular Malaysia, 63(2), 75–80.