Scientia Africana, Vol. 21 (No. 2), August, 2022. Pp 39-52 *https://dx.doi.org/10.4314/sa.v21i2.5*

ISSN 1118 – 1931

APPLICATION OF HIDDEN MARKOV MODEL IN YAM YIELD FORECASTING

Lawal Adamu[1*], Saidu Daudu Yakubu[2],Didigwu Ndidiamaka Edith[3], Abdullahi Abubakar[4], Khadeejah James Audu[1], and Isaac Adaji[5]

[1] Department of Mathematics, Federal University of Technology, Minna, Nigeria
[2]Department of Mathematical Science, Ibrahim Badamasi Babagida University, Lapai, Nigeria
[3]Department of Mathematics, Enugu State University of Science and Technology, Nigeria
[4]Department of Mathematics, Faculty of Science, Air Force Institute of Technology, Kaduna, Nigeria
[5]Department of Mathematics and Stattistics, Confluence University of Science and Technology, Osara, Kogi State.
*lawal.adamu@futminna.edu.ng

Received: 19-02-2022
Accepted: 06-06-2022

ABSTRACT

Providing the government and farmers with reliable and dependable information about crop yields before each growing season begins is the thrust of this research. A four-state stochastic model was formulated using the principle of Markov, each state of the model has three possible observations. The model is designed to make a forecast of yam yield in the next and subsequent growing seasons given the yam yield in the present growing season. The parameters of the model were estimated from the yam yield data of Niger state, Nigeria for the period of sixteen years(2001-2016). After which, the model was trained using Baum-Welch algorithm to attend maximum likelihood. A short time validity test conduct on the model showed good performance. Both the validity test and the future forecast shows prevalence of High yam yield, this attest to the reality on the ground, that Niger State is one of the largest producers of yam in Nigeria. The general performance of the model, showed that it is reliable therefore, the results from the model could serve as a guide to the yam farmers and the government to plan strategies for high yam production in the region.

Keywords: Yam Yield, Hidden Markov Model, Rainfall, Temperature, Transition Probability, Observation Probability

INTRODUCTION

After air and water, the most important thing for continual existence of life on earth is food, hence, providing its security is essential for population growth and development of any nation (Lawal, *et al* 2016). The crop yield prediction is one of the most desirable yet challenging tasks for every nation. Nowadays, due to the unpredictable climatic changes, farmers are struggling to obtain a good amount of yield from the crops (Khosla *et al.*, 2020). At the present time, one of the most important sectors of Nigeria economy is agriculture. It is also the major means of livelihood of many homes in the country. With population of about Two hundred million people, more than 70% of the population is engaged in agricultural activities. To feed the increasing population of Nigeria, there is a need to incorporate the latest technology and tools in the agricultural sector. Accurate and reliable seasonal forecasts of crop yields are among the most valuable pieces of information that stakeholders such as farmers, commodity traders, and government officials can have at their disposal to make strategic decisions in their respective roles (Basso and Lin, 2018). The aim of this research is to provide the government and farmers with a precise, scientific sound and independent forecasts of yam's yield before the growing season begins in the study area by considering the effect of major climatic elements(rainfall and temperature).The study area of this research is Niger State. Niger State has Guinea savannah vegetation which covers the entire landscape and is characterized by woodlands and tall grasses interspersed with tall dense species. It has tropical climate.

Niger state is located on latitude 3.20^{o}East and longitude 11.30°North in the North Central Zone of Nigeria with land mass of 92,800 km^2 and having a population of 3,950,249 with about 85% of this population practicing agriculture specifically, growing yam, rice, sorghum, cowpea and maize in large quantity. Niger Sate is an agriculture-based state in Nigeria. Nigeria produces about 70% of the world's yam accounting for about 39.9 million tons (Zhag *et al.,* 2017). Niger State is one of the largest yam producing State in Nigeria. Rainfall is the most important natural factor that determines the agricultural production in Niger State.

Many researchers within the Nigeria and around the world had use several mathematical methods to forecast crop yield for different crops before or during growing season, among the researchers are: Dahikar and Rode (2014) studied basic requirements for applications of Artificial Neural Networks (ANNs) in yield prediction. Simple network architectures, with one hidden layer and back propagation of errors were tested for different predictors and crops, like cotton, sugarcane, wheat, rice and others. Reported in (Andrew *et al.,* 2007) is a Non-homogeneous Hidden Markov Model (NHMM), the model was used to make stochastic simulations of March–August daily rainfall at 10 stations over the south eastern United States, 1923–1998. Station-averaged observed daily rainfall amount was prescribed as an input to the NHMM, which was then used to disaggregate the rainfall in space. These rainfall simulations were then used as inputs to a Crop Estimation through Resource and Environment Synthesis (CERES) crop model for maize. The regionally averaged yields derived from the NHMM rainfall simulations were found to correlate very high with those generated by the crop model using observed rainfall. (Saran *et al.,* 2006) proposed a new approach to estimate rice cultivation and harvest dates by using 8-day composite normalized difference vegetation index (NDVI) derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data. In their work, they divided the rice growth states into 4 states, namely, nothing, growing, mature, and harvest states and applied them to the Hidden Markov Model (HMM). Then, they assigned the state to the NDVI time-series data by using the Viterbi algorithm. By using those derived states, they were able to estimate the rice cultivation and harvest dates. The date estimation results were compared with the ground truth data to access the accuracy and they found that the average cultivation dates and harvest dates to have errors of 15.48 days and 6.525 days, respectively.

Lawal *et al.,* (2021) analyzed weekly rainfall pattern of Makurdi, Benue State, Nigeria using the principle of Markov, they found that that in the long-run 22% of the weeks during rainy season in Makurdi, will experienced No rainfall, 50% will experienced Low rainfall, 25% will experienced Moderate rainfall and 2% will experienced High rainfall. They also found that, a week of High rainfall cannot be followed by another week of High rainfall, a week of High rainfall cannot be followed by a week of No rainfall, and a week of Moderate rainfall cannot precede a week of High rainfall. Nguyet and Dung (2021) had established a multi-step procedure for using HMM to select stocks from the global stock market. Firstly,

Scientia Africana, Vol. 21 (No. 2), August, 2022. Pp 39-52 https://dx.doi.org/10.4314/sa.v21i2.5
© Faculty of Science, University of Port Harcourt, Printed in Nigeria ISSN 1118 – 1931

they identified and scored five important factors of a stock based on its historical performances. Secondly, HMM was used to predict the regimes of six global economic indicators and find the time periods in the past during which these indicators have a combination of regimes that is similar to those predicted. Successful application of Hidden Markov Model have also been reported in the following researches; reinforcement learning algorithm (Silver *at al*., 2018), identification and inverse filtering (Robert Mattila, 2020), portfolio optimization (Jerndal and Krödel, 2018), in rainfall pattern prediction (Lawal, 2018).

MATERIALS AND METHODS

Study Area and Data source

The study area of this research is Niger State, is one of the states in the middle belt of Nigeria. The data used in this research were collected from the archive of Niger state Bureau of Statistics, for the period of 16 years (2001-2016).

Hidden Markov Model

The mathematical method adopted in this research work is called Hidden Markov Model (HMM). A 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 $s_i \in S$, i=1, N, which are hidden, i.e. not directly observable (Enza and Daniele, 2007)

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 $Q = q_{1,} q_2, q_3, ... q_T$
O = observation sequence
$O = o_{1,} o_{2,} o_3 .... o_T$ $A = \{a_{ij}\}$
Transition probability matrix
Observation probability matrix $B = \{b_j(o_t)\}$
Where $b_j(o_t) = p(o_t \mid q_t = s_j)$
If the observation is continuous a probability density function is used

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

(1)

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

Model Formulation

The amount of crop yield in a growing season depends on several factors, namely genetic, climatic, biotic, edaphic, socio-economic and physiographic factors. However some of these factors have less effect on crop growth and development but there are some factors that are very essential to the growth and development of plant. The most important factor is the climatic factor in which all other factors depend on it. Nearly 50 % of crop yield is attributed to the influence of climatic factors (Lawal *et al*., 2016). The climatic factors are rainfall, temperature, atmospheric humidity, solar radiation, wind velocity. These climatic factors influence all plant growth processes such as photosynthesis, respiration, transpiration, breaking of seed dormancy, seed germination, protein synthesis, and translocation (Ben, 2011). The most important factors for crop production that cannot be neglected or compromised are rainfall and temperature according to Niger state Agricultural and Mechanization Development Authority. Agronomists and soil scientists are interested in precipitation and rainfall in particular as a source of soil moisture to crops. Soil moisture is critical to soil chemical processes. Particularly nitrogen-fixation depends on water availability in the soil and hence soil moisture can contribute substantially to the availability of nutrients. Water supply is usually the most critical factor determining yield. The effects of water shortages on production may vary according to the particular crop, the soil characteristics, the root system, and the severity and timing of shortages during the growth cycle (Ahn, 1993). Most plant processes related to growth and yield are highly temperature dependent. The optimum growth temperature frequently corresponds to the optimum temperature for photosynthesis. Temperature increase can have both positive and negative effects on crop yields. Higher temperature also affects the rate of plant development (vegetative growth) and hence speeds annual crops through the developmental process. Temperature increases, however, have also been found to reduce the yields and quality of many crops, particularly cereal and feed grains. For example, higher temperatures shorten the life cycle of grain crops, resulting in a shorter grain filling period, so the plants produce smaller and lighter grains, culminating in lower crop yields and perhaps poorer grain quality, i.e. lower protein levels (Adams *et al*., 1998; Zhag *et al.*, 2017). Waziri *et al.*, (2014) further pointed out that temperature is very important for accumulation of organic matter as well as ripening of the plants. These two major climatic elements have been considered to model yam yield using HMM. The model is developed to make forecast of yam yield with respect to these climatic elements. Markov models have become veritable tools in crop yields forecast and modelling of other stochastic processes. We consider amount of temperature and rainfall in modelling yam yield using hidden Markov model. The amount of Yam yield largely depend on these climatic elements, they contribute immensely to yam yield in a grown season. But we cannot ordinarily measure how much of each of them contribute to the overall output (the amount of yam yield). The amount of yam yield depends on them and their amounts are not static or deterministic but they varies randomly from year to year, this makes the amount of yam yields in each growing seasons to varies. This situation follow a doubly stochastic process with the amount of yam yield per hectare as the observation of the HMM depending on the state (amount of rainfall and temperature). In view of this, we considered the amount of yam yield per hectare within a growing as an emission of the HMM while the amount rainfall and Temperature within the same period is taking as state of our model, as a result we make the following assumptions.

(i) The transition between the states is governed by first order Markov dependence, as represented by equation (2)

$$P\{X_{t+1}=j \mid X_0=i_0,\ldots,X_{t-2}=i_{t-2},X_{t-1}=i_{t-1},X_t=i\}=P\{X_{t+1}=j \mid X_t=i\}=P_{ij}(t) \qquad (2)$$

(ii) The probability of distribution of generating current observation symbol depends only on current state. That is $P(O \mid Q,\lambda)=\prod_{t=1}^{T}P(o_t \mid q_t,\lambda)$ (3)

(iii) Amount of rainfall is considered to be low if it is below 1203mm

Scientia Africana, Vol. 21 (No. 2), August, 2022. Pp 39-52 *https://dx.doi.org/10.4314/sa.v21i2.5*
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

(iv) Amount of rainfall is considered to be high if it is above 1203mm
(v) Amount of temperature is considered to be low if it is below 30.4^0C
(vi) Amount of temperature is considered to high if it is above 30.4^0C
(vii) Amount of Yam yield is considered to low if it is below 12.2 metric tons
(viii) Amount of Yam yield is considered to moderate if it is within the range (12.2 -15.27) metric tons
(ix) Amount of Yam yield is considered to high if it is above 15.27 metric tons

Let the hidden Markov model for Yam yield forecast be modelled by four states and three observations thus:

State1: Low rainfall and low temperature
State2: Low rainfall and high temperature
State3: high rainfall and low temperature
State4: high rainfall and high temperature
Observations
L = O_1 = Low yield
M = O_2 = Moderate yield
H = O_3= High yield

The classification of states and the observations, and the assumptions made in this work is based on the study area and data collected from the archive of Niger state Bureau of Statistics, for the period of 16 years (2001-2016)

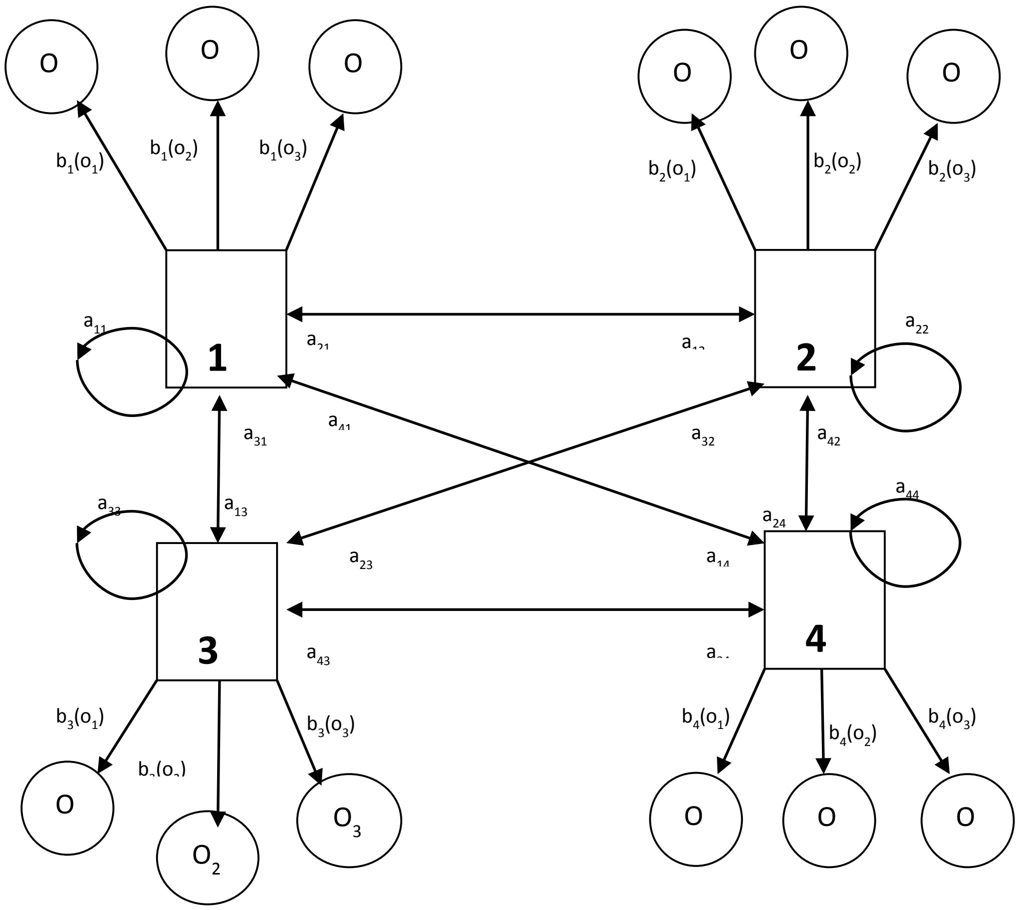

Figure 1: Transition Diagram o f the Model

Transition Probability Matrix for the States

The transition between the states are represented by the matrix below

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \qquad (4)$$

Observation Probability Matrix

The matrix below represents observations emitted from the model

$$B = \begin{bmatrix} b_1(o_1) & b_1(o_2) & b_1(o_3) \\ b_2(o_1) & b_2(o_2) & b_2(o_3) \\ b_3(o_1) & b_3(o_2) & b_3(o_3) \\ b_4(o_1) & b_4(o_2) & b_4(o_3) \end{bmatrix} \qquad (5)$$

Initial Probability Distribution

The initial probability distribution for the model is given below

$$\pi = [\pi_1 \quad \pi_2 \quad \pi_3] \qquad (6)$$

The Hidden Markov Model for Yam Yield Forecast

The overall hidden Markov model for Yam yield forecast is given by the compact notation below

$$\lambda = (A, B\ ,\pi) \qquad (7)$$

Where A is the Transition probability Matrix, B is the observation probability Matrix and π is the initial probability distribution

Method of Forecasting

Likelihood based prediction is adopted in this work and is done along with training of the parameters of the model (Rabiner, 1989). In this method, the parameters of the model are initialized then trained using Baum-Welch algorithm (lawal, *et al* 2016) to attends Maximum likelihood. The forward probability of the training observation sequence is calculated from time t=1 to T using Forward Algorithm in (lawal, *et al* 2016). To forecast the next state at time T+1 and its observation given the present state at time T, forward probability for each possible observations of the states are calculated, the sequence with highest value of the forward probability at time T+1 is taken as forecasted observation and its state. The forecast is made for the next years (time T+2 to time T+n). To avoid underflow of the forward algorithm we let the coefficient

$$c_t = \frac{1}{\sum_{t=1}^{N} \alpha_t(i)} \qquad (8)$$

And thus the new scaled value for α becomes

Scientia Africana, Vol. 21 (No. 2), August, 2022. Pp 39-52 https://dx.doi.org/10.4314/sa.v21i2.5
© Faculty of Science, University of Port Harcourt, Printed in Nigeria ISSN 1118 – 1931

$$\hat{\alpha}_t(i) = c_t \times \alpha_t(i) = \frac{\alpha_t(i)}{\sum_{t=1}^{N} \alpha_t(i)} \qquad (9)$$

RESULTS AND DISCUSSION

The data used in this illustration was collected from the archive of Niger state Bureau of Statistics for the period of 16 years (2001-2016). Yam is the major cash crop in Niger state. Niger state is an agriculture-based state in Nigeria. Nigeria produces about 70% of the world's yam accounting for about 39.9 million tons (Waziri *et al*., 2014). Niger State is one of the largest yam producing states in Nigeria. Rainfall is the most important natural factor that determines the agricultural production in Niger State. The summary of states and observations distribution obtained from the raw data is presented in Table 1

Table 1: Summary of State and Observations Distribution of Yam Yield for a Period of Thirteen Years

Year	State	Observation
2001	4	L
2002	2	L
2003	4	L
2004	1	L
2005	2	M
2006	4	H
2007	3	H
2008	3	L
2009	4	H
2010	2	M
2011	2	H
2012	2	H
2013	4	M
2014	4	H
2015	4	H
2016	4	H

Validity Test for the Model

To test for the validity of the model, we estimate the parameters of HMM1 using the rainfall , temperature and yam yield data from 2001 to 2013, then make forecast for 2014, 2015 and 2016

Transition Count Matrix

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 0 \end{bmatrix} \qquad (10)$$

Pseudocount Transition Matrix

Pseudocount is added to the transition count matrix to avoid zero probability during the training this may be a problem to the mechine/computer (Srinivas, 2006). Thus

$$T = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & 1 & 4 \\ 1 & 1 & 2 & 2 \\ 2 & 3 & 2 & 1 \end{bmatrix} \qquad (11)$$

Transition Probability Matrix

$$A = \begin{bmatrix} 0.2000 & 0.40000 & 0.2000 & 0.2000 \\ 0.1111 & 0.3333 & 0.1111 & 0.4444 \\ 0.1666 & 0.1666 & 0.3333 & 0.3333 \\ 0.2500 & 0.3750 & 0.2500 & 0.1250 \end{bmatrix} \qquad (12)$$

Observation Count Matrix

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix} \qquad (13)$$

Pseudocount Observation Matrix

$$E = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 3 \end{bmatrix} \quad (14)$$

Observation Probability Matrix

$$B = \begin{bmatrix} 0.2222 & 0.1428 & 0.1111 \\ 0.2222 & 0.4290 & 0.3333 \\ 0.2222 & 0.1428 & 0.2222 \\ 0.3333 & 0.2850 & 0.3333 \end{bmatrix} \quad (15)$$

Initial State Probabilities

$$\pi = \begin{bmatrix} 0.0769 & 0.3846 & 0.1538 & 0.3846 \end{bmatrix} \quad (16)$$

$$\lambda_1 = (A, B, \pi) \quad (17)$$

Equation (17) is the test HMM1

After 1000 iterations of the Baum Welch Algorithm using MatLab 2015, equation (17) converged to equation (21)

$$\hat{A} = \begin{bmatrix} 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 0.0242 & 0.4000 & 0.0000 & 0.5758 \\ 0.2500 & 0.0000 & 0.7500 & 0.0000 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 \end{bmatrix} \quad (18)$$

$$\hat{B} = \begin{bmatrix} 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 \\ 1.0000 & 0.0000 & 0.0000 \\ 0.3473 & 0.6527 & 0.0000 \end{bmatrix} \quad (19)$$

$$\hat{\pi} = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \quad (20)$$

$$\lambda_1^* = (\hat{A}, \hat{B}, \hat{\pi}) \quad (21)$$

Making Forecast for 2014

From our dataset (Table 1), the process is in State 4 at time T (2013) emitting Observation M. To get the next observation at time T+1, we calculate the forward probability for ML, MM, and MH, then take the one with highest liklihood value in Table 2

Table 2: Likelihood Based Prediction Table for 2014 Yam yield Forecast

	ML	MM	MH
STATE 1	00	0.0	0.0
STATE 2	0.0	0.0	1.0
STATE 3	0.0	0.0	0.0
STATE 4	0.0	0.0	0.0

From table 2, State 2 under MH, has the highest likelihood value, so is taken as most probable observation sequence at T+ 1, (that is, State 2 emitting observation H at T+1)

States: **4**(2013) $\rightarrow$ **2**(2014)

$\downarrow$ $\downarrow$

Scientia Africana, Vol. 21 (No. 2), August, 2022. Pp 39-52 *https://dx.doi.org/10.4314/sa.v21i2.5*
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

Observations: M H

From the computation of Table 2, the process is in state 2 at time T+1(2014) with observation H and sequence MH. To get next sequence at time T+2, we calculate the forward probability for MHL, MHM, and MHH and take the one with the highest likelihood value in Table 3

Table 3: Likelihood Based Prediction Table for 2015 Yam Yield Forecast

	MHL	MHM	MHH
STATE 1	0.0	0.0242	0.0
STATE 2	0.0	0.0	0.4
STATE 3	0.0	0.0	0.0
STATE 4	0.1999	0.3758	0.0

From table 3, State2 under MHH has the highest likelihood value, so is taken as most probable observation sequence at T+ 2 (that is, State 2 emitting observation H at T+2)

States: **4**(2010) → **2**(2014) → **2**(2015)

↓ ↓ ↓

Observations: M H H

From the computation of Table 3, the process is in state2 at tme T+2(2015) with observation H and sequence MHH. To get the next sequence at time T+3, we calculate the forward probability for MHHL, MHHM, and MHHH and take the one with the highest likelihood value in Table 4

Table 4: Likelihood Based Prediction Table for 2016 Yam Yield Forecast

	MHHL	MHHM	MHHH
STATE 1	0.0	0.024	0.0
STATE 2	0.0	0.0	0.4
STATE 3	0.0	0.0	0.0
STATE 4	0.1998	0.3759	0.0

From table 4, State 2 under MHHH has the highest likelihood value, so is taken as most probable observation sequence at T+ 3, (that is, State 2 emitting observation H at T+3)

States: **4**(2013) → **2**(2014) → 2(2015) →2(2016)

↓ ↓ ↓ ↓

Observations: M H H H

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

Predicted States and observations

$$\begin{array}{lcccc} \text{States:} & \mathbf{4}(2013) \rightarrow & \mathbf{2}(2014) \rightarrow & 2(2015) \rightarrow & 2(2016) \\ & \downarrow & \downarrow & \downarrow & \downarrow \\ \text{Observations:} & M & H & H & H \end{array}$$

Actual states and observations

$$\begin{array}{lcccc} \text{States:} & \mathbf{4}(2013) \rightarrow & \mathbf{4}(2014) \rightarrow & 2(2015) \rightarrow & 2(2016) \\ & \downarrow & \downarrow & \downarrow & \downarrow \\ \text{Observations:} & M & H & H & H \end{array}$$

Hidden Markov Model(HMM2) for future forecast

HMM2 is developed to forecast yam yield for future years, the parameters of the HMM2 was estimated using the rainfall, temperature data from 2001 to 2016, then make forecast for 2017, 2016, 2017, 2018, 2019 and 2020. Thus we have

Transition Count Matrix

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 3 \end{bmatrix} \tag{22}$$

Pseudocount Transition Matrix

$$T = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & 1 & 4 \\ 1 & 1 & 2 & 2 \\ 2 & 3 & 2 & 4 \end{bmatrix} \tag{23}$$

Transition Probability Matrix

$$A = \begin{bmatrix} 0.2000 & 0.40000 & 0.2000 & 0.2000 \\ 0.1111 & 0.3333 & 0.1111 & 0.4444 \\ 0.1666 & 0.1666 & 0.3333 & 0.3333 \\ 0.1818 & 0.2727 & 0.1818 & 0.3636 \end{bmatrix} \tag{24}$$

Scientia Africana, Vol. 21 (No. 2), August, 2022. Pp 39-52 https://dx.doi.org/10.4314/sa.v21i2.5
© Faculty of Science, University of Port Harcourt, Printed in Nigeria ISSN 1118 – 1931

Observation Count Matrix

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 5 \end{bmatrix} \tag{25}$$

Pseudocount Observation Matrix

$$E = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 6 \end{bmatrix} \tag{26}$$

Observation Probability Matrix

$$B = \begin{bmatrix} 0.2222 & 0.1429 & 0.0833 \\ 0.2222 & 0.4285 & 0.2500 \\ 0.2222 & 0.1428 & 0.1666 \\ 0.3333 & 0.2857 & 0.5000 \end{bmatrix} \tag{27}$$

Initial state probabilities

$$\pi = \begin{bmatrix} 0.0625 & 0.3125 & 0.1250 & 0.5000 \end{bmatrix} \tag{28}$$

$$\lambda_2 = (\pi, A, B) \tag{29}$$

After 1000 iterations of the Baum Welch Algorithm, equation (29) converged to equation (33)

$$\hat{A} = \begin{bmatrix} 0.5000 & 0.5000 & 0.0000 & 0.0000 \\ 0.000 & 0.000 & 0.0000 & 1.0000 \\ 0.5000 & 0.0000 & 0.5000 & 0.0000 \\ 0.0000 & 0.4286 & 0.0000 & 0.5714 \end{bmatrix} \tag{30}$$

$$\hat{B} = \begin{bmatrix} 1.0000 & 0.0000 & 0.0000 \\ 0.2500 & 0.7500 & 0.0000 \\ 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \tag{31}$$

$$\hat{\pi} = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \tag{32}$$

$$\lambda_2^* = (\hat{A}, \hat{B}, \hat{\pi}) \tag{33}$$

Forecast for 2016 to 2020

Following similar method of prediction procedures and using equation(30) to (32)the results for 2017 and 2020 is presented below

$$\begin{array}{llllll} \text{States:} & \mathbf{4}(2016) \rightarrow & \mathbf{4}(2017) \rightarrow & \mathbf{4}(2018) \rightarrow & \mathbf{4}(2019) \rightarrow & \mathbf{4}(2020) \\ & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ \text{Observations:} & H & H & H & H & H \end{array}$$

The parameters of the HMM1 was estimated using rainfall, temperature and yam yield data from 2001 to 2013, after 1000 iterations of the Baum algorithm, λ_1 converged to a new model, λ_1^* this new model was then used to forecast yam yield for 2014, 2015 and 2016. From the forecast time series the HMM1 was in state 4 at time **T**(2013) emitted Moderate yield, then make transition to state 2 at time T+1(2014) emitting High yield. Similar interpretation is given to transition to state 2 at **T+2**(2015) and transition to state 2 at time **T+3** (2016) both emitting High yields. The transitions between the states are govern by the first order Markov dependence as mentioned in the previous sections. This short time validity test shows 100% accuracy in the yam yield forecast when compared with the yam yield from the dataset.

The parameters of the HMM2 was estimated using rainfall, temperature and yam yield data from 2001 to 2016, after 1000 iterations of the Baum algorithm, λ_2 converged to a new model, λ_2^*, this new model was then used to make forecast for future years . From the forecast time series the HMM2 was in state4 at time **T** (2016) emitted High yield, then make transition to state **4** at time **T+1** (2017) emitting High yield. Similar interpretation is given to transition to state **4** at **T+2** (2018), transition to state 4 at time **T+3** (2019) and transition to state 4 at time T+4(2020) all emitting High yields.

CONCLUSION

In this paper, a hidden Markov model to forecast yam yield with respect to weather parameters has been developed and implemented in Niger state, Nigeria. The short time validity test for the model showed good performance. Both the validity test and the future forecast shows prevalence of High yam yield, this attest to the reality on the ground, that Niger State is one of the largest producers of yam in Nigeria. The results from the validity test, shows that the model is reliable and dependable. Therefore, the results from the model could serve as a guide to the farmers and the government to plan strategies for high yam production in Niger state. The model could also be used to forecast yields of other crops with little or no modifications.

REFERENCES

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.

Scientia Africana, Vol. 21 (No. 2), August, 2022. Pp 39-52 https://dx.doi.org/10.4314/sa.v21i2.5
© Faculty of Science, University of Port Harcourt, Printed in Nigeria ISSN 1118 – 1931

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

Andrew W. R, Amor V. M., & James W. H. (2007). Downscaling of Seasonal Precipitation for Crop Simulation, *Journal of Applied Meteorology and Climatology* vol 46 p677-693

Basso and, Lin Liu(2018) Seasonal crop yield forecast: Methods, applications, and accuracies. Advances in Agronomy, Volume 154 pp201-245.

Ben G. Bareja (2011). Cropreview.com .http://www.cropsreview.com/climaticfactors.htmlAccesed 20/8/2015

Dahikar MSS and Rode SV (2014). Agricultural crop yield prediction using artificial neural network approach. International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering (IJIREEICE). 2014 Jan; 2(1):1–4.

Enza Messina, Daniele Toscani (2007). Hidden Markov models for scenario generation, IMA *Journal of Management Mathematics* 19: 379-401

Jerndal. R and O. Krödel (2018), “Portfolio optimization with market state analysis,” Bachelor’s thesis, KTH Royal Institute of Technology,

Khosla, E., Dharavath, R. & Priya, R(2020). Crop yield prediction using aggregated rainfall-based modular artificial neural networks and support vector regression. *Environ Dev Sustain* **22,** 5687–5708 https://doi.org/10.1007/s10668-019-00445-x

Lawal Adamu (2018). “Hidden Markov Model and its Application in Rainfall Pattern Prediction” (A Monograph in Mathematics) LAP Lambert Academic Publishing: ISBN: 978-613-4-98532-1

Lawal, A; Abubakar, UY; Danladi, H; Gana, AS(2016) Prediction of Annual Rainfall Pattern Using Hidden Markov Model (HMM) in Jos, Plateau State, Nigeria: *Journal of Applied Sciences and Environmental Management*

Lawal. Adamu, U.Y. Abubakar, D.Hakimi and A.S. Gana(2021) Prediction of Weekly Rainfall Both in Discrete and Continuous Time Using Markov Model (Chapter 9) *International Research in Environment, Geography and Earth Science Vol. 8.* Book publisher International. 127-143

Nguyet Nguyen and Dung Nguyen (2021).Global Stock Selection with Hidden Markov Model. Risks 9: 9. https://doi.org/ 10.3390/risks9010009

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

Robert Mattila (2020). Hidden Markov Models: Identification, Inverse Filtering and Applications Doctoral Thesis in Electrical Engineering KTH Royal Institute of Technology Stockholm, Sweden

Saran S, Teerasit K, Kitti C, Preesan R, Thitiporn C, Itsuo K. (2014). Rice Cultivation an Harvest Date Estimation Using MODIS NDVI Time-series Data R2-59, 4266 Nagatsuta-cho, Midori-ku, Yokohama, 226-8503, Japan

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

Syngenta (2018). Syngenta Crop Challenge In Analytics. Available online at: https://www.ideaconnection.com/syngenta-crop-challenge/challenge.php/

Waziri A., Tsado E.K., Likita .T , Gana, A.S.(2014).Socio-economic factors influencing adoption of yam Minisett Technology in Niger State of Nigeria *Journal of Biology, Agriculture and Healthcare* 4(5), 98-105