Volume 24(3), 58 - 64, 2020
JOURNAL of Horticulture, Forestry and Biotechnology
www.journal-hfb.usab-tm.ro

GARCH Model Analysis of Harvested and Healthy Cocoa Pods

Dauda Taofik. Oyedele*, Fasade, A. A., and Olasehinde, L.O.

Institute of Agricultural Research and Training, Obafemi Awolowo University, PMB 5029, Moor Plantation, Ibadan, Nigeria.

*Corresponding author email: todauda.iart@gmail.com

Abstract Volatility clustering, leptokutorsis and leverage effects are main features of time series data and financial markets data likewise production data often exhibit volatility clustering where time series show periods of both high and low volatility. This study was conducted to model the conditional mean and variance of cocoa production and number of healthy cocoa in South West Nigeria using total cocoa pods harvested (THP) and Number of healthy pods (NHO) in SW Nigeria (1991 – 2011). The autocorrelation, r_k of the two commodities over a convenient lag of 77 ((N/2) – 1)) showed that the autocorrelation was non-random because 0.554 (<0.95) of the r_k falls within ±2/√156 = ± 0.16013. The 95% confidence interval of the r_k indicated that the number of healthy pod has a shorter autocorrelation length (47) than the total harvested pods (75). The $GARCH_{(1,1)}$ model coefficient returned persistence parameters of 0.85955 (For THP) and 0.860754 (for NHO) and these were less than unity. These indicated that the GARCH models have unique and stationary solution. Also, the half-life of the GARCH models for THP was 4.57987 while that of NHO was 4.62263hence, the $GARCH_{(1,1)}$ model could be said to be persistent since their 1/2 – life were finitude. With the GARCH (1,1) model, the risk value of production process can be estimated and such risks could also be hedged out.

Key words

leptokutorsis, leverage, risk, clustering, volatility, hedge

Let y_t be a random value of productivity (of cocoa) at a time t, the auto-regressive values, y_t can be determined using;

$$AR_{(1)} = y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon_t \quad .(1)$$

Where y_t is a random variable drawn from a conditional density function $f\{y_t|y_{t-1}\}$ and ε_t is white noise [1]. The above equation has mean and the variance $V(\varepsilon_t) = 0$.

From this $AR_{(1)}$model, the forecast of today's value (of cocoa produced) y_t conditioned on past information is simply,

$$E\{y_t \mid y_{t-1}\} = \beta_0 + \beta_1 y_{t-1} \quad .(2)$$

Improvements in the forecasting accuracy of time series models are attributable to the optimal use of past information. The past information was however usually generated amidst instability in several of the variables thus leading to higher variance in the data. The prediction of today's values can thus be improved by the incorporation of the volatility of the variance into the model.

Volatility clustering, leptokutorsis and leverage effects are main features of time series data [2]. Financial markets data often exhibit volatility clustering where time series show periods of both high and low volatility. Indeed, with economic and financial data, time-varying volatility is more common than constant volatility and accurate modeling of time-varying volatility is of great importance in financial engineering [3]. The peculiarities of volatility can however be found in production data too and thus given the possibilities of its applicability in production phenomena. Auto-regressive moving average (ARMA) models are used to model the conditional expectation of a process given the past. However, in ARMA model, the conditional variance given the past is constant. The inability of ARMA model to capture this type of behavior provide the justification for GARCH time series models and it is becoming widely used in econometrics and finance because they have randomly varying volatility processes [3]. AutoRegressive Conditional Heteroscedasticity (ARCH) models the conditional variance and has a structure very similar to the structure of the conditional expectation in an AR model.

Autoregressive conditional heteroskedasticity (ARCH) model is the first model of conditional Heteroskedasticity ([4] and [5]) and it is given as;

$$y_t = x_t \xi + \varepsilon_t \qquad t = 1, \ldots, T.$$

Where x_t = K x 1 vector of exogenous variable, ξ = k x 1 vector of regression parameters, ε_t = stochastic error.

Similarly, x_t has conditional variance given by

$$x_t^2 = \alpha_0 + \alpha_1 \mu^2_{t-1}$$

Literatures on GARCH are colossal and include, [5], [6], [7] and [8]. Some of these literatures focused on volatility of commodities' prices and inflation. Little or scanty works on volatility of production of goods are known. The purpose of using Cocoa for this study was to analyze systematic mistrending in the productivity of commodity in the SW Nigeria since it would provide the nature of volatility in the cocoa productivity. The goal of this study was thus application of GARCH model to forecast the volatility of cocoa productivity in the South-West Nigeria. Similarly, pricing is an economic activity just like production that are liable to fluctuation that can bring about volatility hence the basis of this study was to investigate the adoptability of the tools in cocoa productivity. Plausible model for predicting volatility in cocoa productivity provides a support for strategic management of risk associated with cocoa cultivation. The objective of the study was therefore to model the conditional mean and variance of cocoa production and number of healthy cocoa in SW Nigeria.

Materials and Methods

The data was obtained from Cocoa Research Institute of Nigeria (CRIN). It composed of total annual cocoa pods harvested (THP) and Number of healthy pods (NHO) in SW Nigeria over a convenient period of 12 years (1991 – 2011). Nigeria is one of the largest producers of cocoa and it (Nigeria) is broadly divided into six geo-political regions. These are North-North, North-Central, North-East, South-South, South-East and South-West. It is located between Latitudes 4^0N - 14^0N and Longitude 2^0E-14^0E. It has a vast land area of 923,768km² (98.6% – land and 1.4% - water). The data obtained were subjected to summary statistics (Mean, variance and standard error of means), time plots, autocorrelation analysis and correlelogram. The autocorrelation was computed over a convenient lag of 77 $(N/2 - 1)$ and their status was evaluated using;

$$\pm 2/\sqrt{N}$$

If 95% of the autocorrelation values, r_k fall within the above region, then r_k are random otherwise, it is either stationary or cyclic. The GARCH models obtained (using Gaussian Error Distribution-GED) were tested for uniqueness, stationary as well as persistence. The t-error distribution for GARCH model returned invalid/error results hence, only GED estimated GARCH was considered in our study. A GARCH model is said to be unique and stationary if;

$$\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1 \text{ and}$$

Persistent if $(\log 0.5 / \log(\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j))$ is finite.

The data were analyzed using both Jmulti and E view (version 8).

Results and Discussions

Summary Statistics and Autocorrelation (R_k) of the Total Harvest and Healthy Cocoa

The descriptive statistics shows that the mean annual production returned for both harvested (THP) and healthy (NHO) cocoa pods were 31426.35 kg and 22608.43 kg (Table 1). Harvested cocoa pods productivity ranged between 0 and 158887 while that of healthy cocoa pods ranged between 0 and 123065.

The implication of this result is that both pods are sufficiently produced in the country with more harvested pods than the healthy cocoa pods since not all harvested cocoa pods would be healthy.

Table 1

Summary Statistics of the Harvested and Healthy Cocoa Pods Productivity

Statistics	Harvested pods	Healthy pods
Minimum	0	0
Maximum	158887	123065
Mean	31426.35	22608.43
SEM	2239.15	1827.51
CV	88.99%	100.96%
Sum of square	$1.21233505714 \times 10^{10}$	$8.0756175764 \times 10^{9}$
Variance	782151411	521000255

The time plot of both total cocoa harvested and the healthy pods showed that the 2 parameters were cyclic and showed a zero-variance trends for the periods (Figure 1A). The change in mean and variance of both THP and NHO showed no systematic trend (Figure 1B). Hence, the time series could be described as stationary time series.

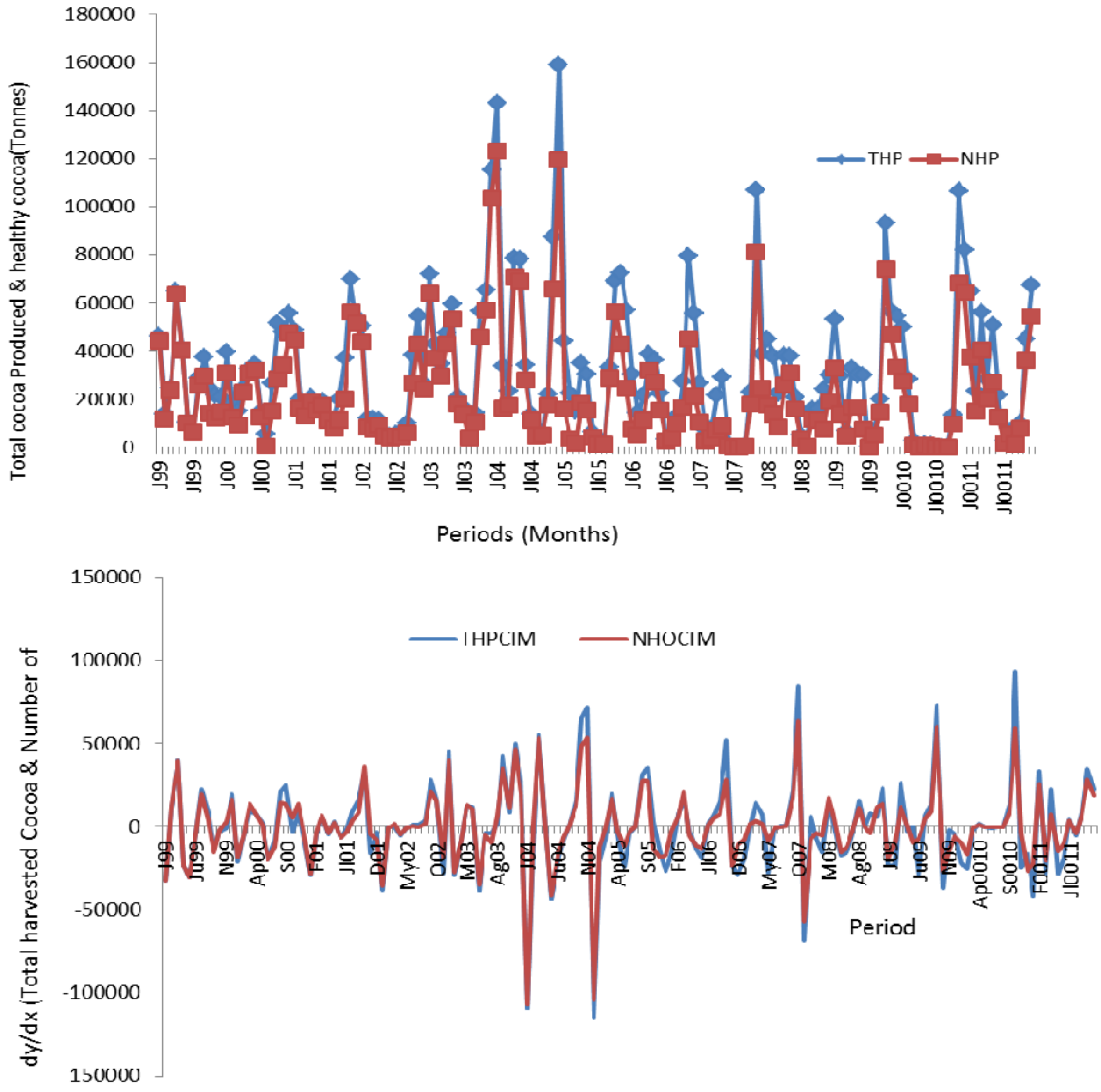


Fig. 1. Time Plot (A) & Change in Mean (B) of the Total cocoa Harvested (THP) and Healthy Pod (NHP)

THP and NHO were established to follow a stationary time series and it might be due to the fact that NHO is a derivative of THP. This was supported by the similarities in the trends of both THP and NHO. The adjusted R^2 of the Cocoa production models (Laderach et al., 2013) were very low and ranged between 0.13 and 0.5. This was not unconnected with the volatility of the cocoa data. It is thus pertinent to adopt the GARCH model which incorporates the volatility of the cocoa productivity. The slow decline of the autocorrelation of the cocoa harvested and healthy cocoa indicated that GARCH(1,1) obtained in this study would be adequate in describing the errors.

The autocorrelation rk of the two commodities ove r a convenient lag of 77 ((N/2) – 1)) showed that the autocorrelation is non-random because 0.554 (<0.95) of the r_k falls within $\pm 2/\sqrt{156} = \pm\ 0.16013$ (Figure 2A). The 95% confidence interval of the autocorrelation values (Figure 2B) showed that the number of healthy pod has a shorter autocorrelation length (47) when compared with autocorrelation length of the total harvested pods (75).

GARCH Model and variance-covariance Matrix of the Model Components.

The GARCH(1,1) model has been defined as;

$$y_k = \sigma_k \varepsilon_k \text{ and}$$

$$\sigma_k^2 = \lambda + \sum_{i=1}^{p} \alpha_i y_{k-1}^2 + \sum_{j=1}^{q} \beta_j \sigma_{k-j}^2 \qquad (3)$$

Where λ = the intercept and > 0, $\alpha, \beta_1 \ \& \ \beta_2$ = last period (t-1) volatility, impact of long term volatility and leverage effects. Volatility in the total cocoa harvested can be estimated using the above GARCH model. The estimated model can be written thus;

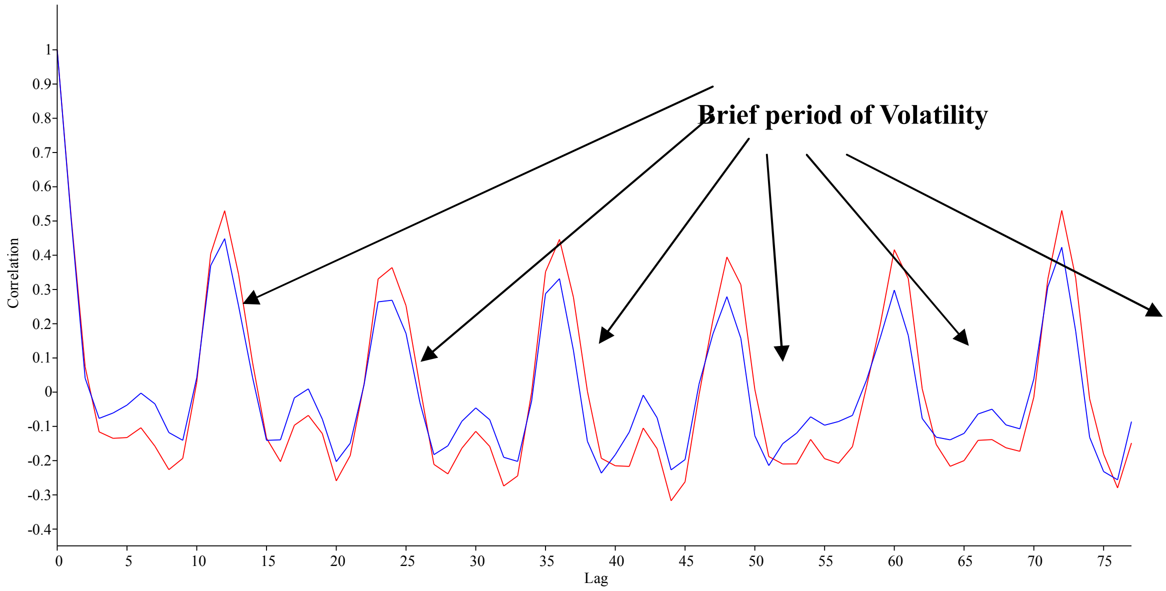
$$y_t = 2.373\times10^{8} + 0.0699\alpha_{t-1} + 0.0005\beta_1 + 0.789\beta_2 \quad (3)$$

$$y_t = 1.334\times10^{8} + 0.0739\alpha_{t-1} + 0.0019\beta_1 + 0.785\beta_2 \quad (4)$$

The former model (3) was the GARCH model for THP while the latter (4) was the GARCH model for the number healthy pods. From the GARCH model coefficient (Table 2), it was obtained that $\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_j$ (persistence parameters) were 0.85955 (For THP) and 0.860754 (for NHO) and these were less than unity. These results indicated that our GARCH models have unique and stationary solution. The half-life of the GARCH models for THP $(\log 0.5/\log(\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_j))$ was 4.57987 while that of NHO was 4.62263 (Table 2). The GARCH model could therefore be said to be persistent since their 1/2 – life were finitude.

The implication of these results is that the impact of shocks on cocoa production or healthy pod's volatility do die away quickly. The asymmetric term in addition, was positive (0.0699 – THP and 0.0739 - NHO) and not significantly different from zero. This implied that negative production shocks or news capable of dampening higher productivity/healthy pods decreases productivity than news capable of inducing productivity/healthy pods.

The variance - covariance matrix of the total harvested cocoa pod ranged between -3.831 x 10^{7} (for the covariance of λ, $_{\beta1}$) and (9.737 x 10^{16} (for the variance of the intercept) while that of number of healthy pods ranged between -1.161 x 10^{7} (for the covariance of λ, $_{\beta1}$) and 1.737 x 10^{16} (for the variance of the intercept – Table 3) The variance – covariance matrix analysis of the GARCH model component revealed that 0.5 of the values were positive while the remaining were negatives. Similar trends was obtained for the Healthy pods and there was no zero variance - covariance values. This implied that, none of the GARCH model components are independent of one another and that as the variance of some of the components affects others positively the variance of some of the components affects others negatively (Table 3).



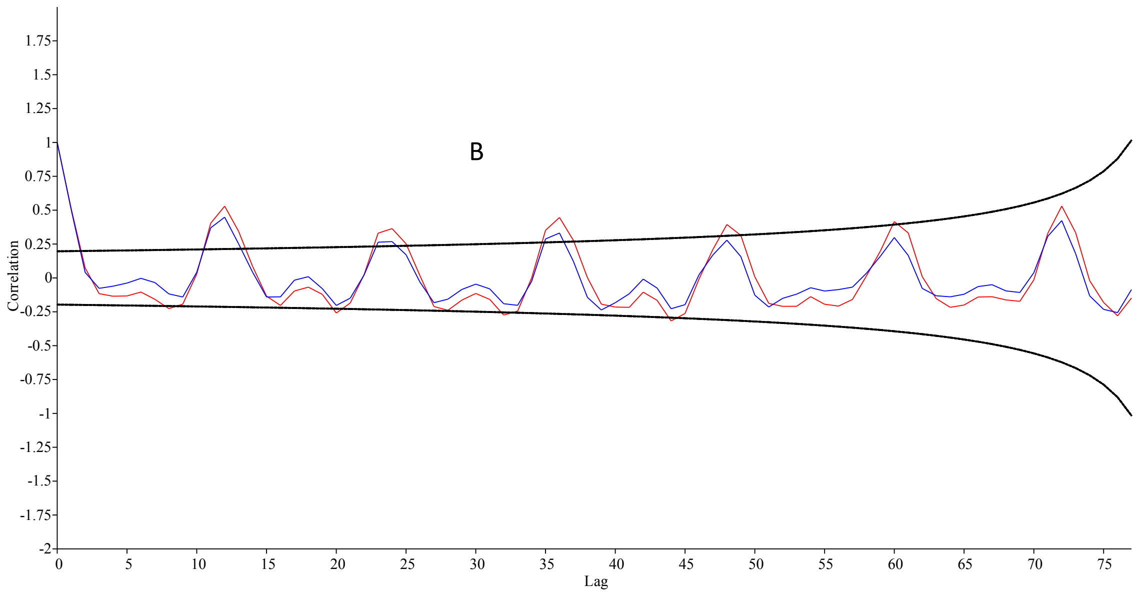


Fig. 2. Autocorrelation of Total cocoa harvested (red line) and Number of Healthy Pods (Blue lines) in tonnes(A) and 95% Confidence interval of the autocorrelation data(B).

Table 2

GARCH Model coefficient and Model Statistics

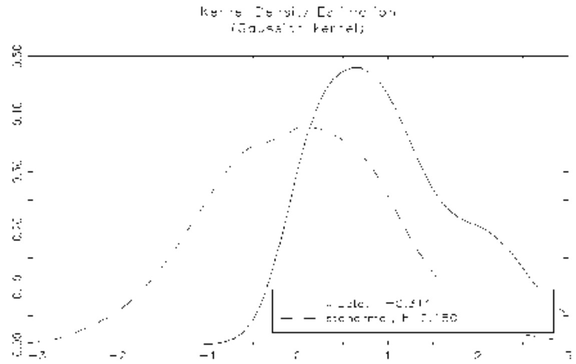
	Variables	**Healthy Pods**		**Harvested Pods**	
		Coefficient	t –statistics	Coefficient	t –statistics
Model Coefficients	Intercept	1.334×10^{8}	1.0140	2.3733×10^{8}	0.76073
	Gamma (1)	0.0739	2.2554	0.06994	1.888
	Beta (1)	0.001954	0.00975	0.0005099	0.00213
	Beta (2)	0.7849	4.22012	0.7891	3.657
Model Statistics	Test (ARCH-LM)	8.261		4.447	
	F - statistics	2.184		1.145	
	Test (Jarque-Bera)	80.0749**		76.0964**	
	Kurtosis	4.9385		5.0517	

Table 3

Covariance Matrix of Harvested Pods and Healthy Pods

		Λ	α_1	β_1	β_2
Harvested Pods	Λ	9.737×10^{16}			
	α_1	5.824×10^{6}	1.373×10^{-3}		
	β_1	-3.831×10^{7}	-2.400×10^{-3}	5.743×10^{-2}	
	β_2	-2.549×10^{7}	-2.421×10^{-3}	-3.036×10^{-2}	4.657×10^{-2}
Healthy Pods	Λ	1.7372×10^{16}	2.600×10^{6}	-1.161×10^{7}	-9.424×10^{6}
	α_1		1.077×10^{-3}	-2.272×10^{-3}	-1.57×10^{-3}
	β_1			4.015×10^{-2}	-2.403×10^{-2}
	β_2				3.459×10^{-2}

The graph of normal distribution showed some improvement in the normalty of the residual (Figure 3A & 3B). The Jarque-Bera diagnosis also indicated that the normalty of the data has been improved. This was because the Jarque-Bera statistics of 80.075 and 76,0964 returned for both THP and NHO were significant ($p < 0.01$ – Table 2). The chi square statistics increased from 4.4467 to 76.0964 for the total harvested pod while the healthy pod increased from 8.2607 to 80.0749. We can therefore choose to reject the null hypothesis of independent distribution of error. The kurtosis returned for the total harvested pod was 4.939 while 5.0577 was obtained for the healthy pods (Table 2). These kurtosis could be described as "leptokurtic" and are expected to have fat tail.



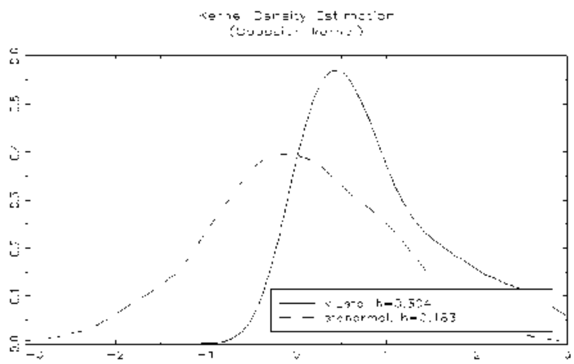


Fig. 3. Model Diagnostic

Lastly, from these results, it is apparent that relationships exist between total cocoa harvested and the number of healthy pods and that this affects subsequent analysis and modelling. The GARCH (1,1) specification arrived at in this study would be enough to model conditional volatility in cocoa productivity. Meanwhile, the GARCH (1,1) model have been established to be a parsimonious representation of conditional variance of many time series data [9] and [10]. Similarly, GARCH (1,1) model was established to outperformed other investigated model in forecasting cocoa bean prices [11]. The positive asymmetry obtained in our study conflict with [12] and this might be due to differences in the study subjects. The object of study in [12] was inflation while cocoa was the object of study in our own study. Our study however agrees in term of positive asymetrism with food price volatility [13] and that of some agricultural crop product in Amhara National Regional State, Ethiopia [14]. The leptokurtic nature of the GARCH (1,1) model obtained in our study fell in line with Gaussian error state of the estimation.

Conclusion

From our empirical results, in conclusion, the cocoa production and health data (THP & NHO) exhibited the main characteristics of time series data which are volatility clustering, leptokurtic and asymmetric effect. The concept of volatility measurement has been widely adopted by econometrician due to the variability nature of the concept/phenomena (environmental and edaphic phenomena) from which data are usually generated. Meanwhile, production data also share some peculiarities of volatility hence the adoption of the tools in our study. It is therefore worthy of note that the essence of volatility measurement in econometrics is sufficient for its adoption in production or other field too. With the GARCH (1,1) model, the value at risk of production process can be estimated and such risks could also be hedged out. The model could also provide guide in optimum allocation of assets.

References

1. Chatfield, C. 1984. The Analysis of Time Series: An Introduction. New York, NY: Chapman and Hall, 268pp.

2. Alberga, D., Shalita, H. and Yosef, R. 2008. Estimating stock market volatility using asymmetric GARCH models. Applied Financial Economics, 18:1201–1208.

3. Engle, R. F. 2011. Long-term skewness and systemic risk. Journal of Financial Econometrics, 9(3):437–468.

4. Taati, S., Hakimipour, N., Alipour, M.S., Saberi, R. and Faramarzi, A. 2015. Analysis of Conditional Asymmetric Volatility of Real GDP and Main Economic Sectors Growth Rates in Iran. IJRRAS 23 (1): 1-14.

5. Bera, A.k. and Higgin, M.L. 1993. ARCH models: Properties Estimation and Testing. Journal of Economic Surveys. 7(4):305-366.

6. Straumman, D. 2005. Estimation in Conditionally Heteroscedastic Time Series Models. Lecture Notes in Statistics, Berlin: Springer – verlag. Volume 181, 365pp

7. Berkes, I., Horvath, L. and Kokoszka, P. 2003. GARCH processes: Structure and Estimation. Bernoulli. 9:201 – 227.

8. Engle, R. 2001. Garch 101: the use of arch/garch models in applied econometrics. *Journal of economic perspectives. 15 (4):157 – 168.*

9. Bollerslev, T. 1987. A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *Review of Economics and Statistics* 69, 542–547.

10. Engle, R. 1993. Statistical Models for Financial Volatility. *Financial Analysts Journal* 49, 72–78

11. Asis, K. Amran, A. and Remali, Y. 2010. Forecasting Cocoa Bean Prices Using Univariate Time Series Model. Journal of Arts Science and Commerce. 1(1): 71 – 80.

12. Omotosho, B.S. and Doguwa, S.I. 2013. Understanding the Dynamics of Inflation Volatility

in Nigeria: A GARCH Perspecctive. CBN Journal of Applied Statistics. 3 (2):51 – 73.
13. Gilbert, C. L. and Morgan, C. W. 2010. Food price volatility. Philosophical Transaction of Royal Society- B 365: 3023–3034
14. Kelkay, B. D. and G/Yohannes, E. 2014. The Application of Garch Family Models to Some Agricultural Crop Products in Amhara National Regional State. Journal of Economics and Sustainable Development. 5 (11):24-35.